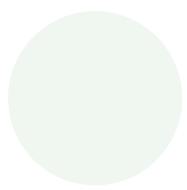
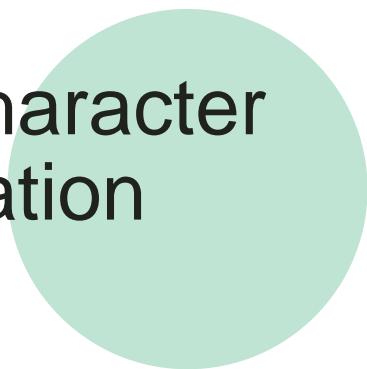
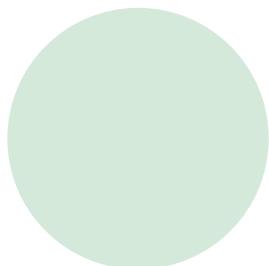
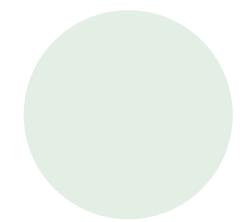


Lessons in Character Impact Evaluation





Lessons in Character Impact Evaluation

Final report

March 2012

Authors:

Dr. Thomas Hanson, Principal Investigator
WestEd

Dr. Barbara Dietsch, Principal Investigator
WestEd

Hong Zheng
WestEd

Project Officer:

OK-Choon Park
Institute of Education Sciences



NCEE 2012-4004
U.S. Department of Education

U.S. Department of Education
Arne Duncan
Secretary

Institute of Education Sciences
John Q. Easton
Director

National Center for Education Evaluation and Regional Assistance
Rebecca A. Maynard
Commissioner

March 2012

This report was prepared for the National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, under contract ED-06C0-0014 with Regional Educational Laboratory West administered by WestEd.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the report.

This report is in the public domain. Authorization to reproduce it in whole or in part is granted. While permission to reprint this publication is not necessary, the citation should read: Hanson, T., Dietsch, B., and Zheng, H. (2012). *Lessons in Character Impact Evaluation*. (NCEE 2012-4004). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

This report is available on the Institute of Education Sciences website at <http://ncee.ed.gov> and the Regional Educational Laboratory Program website at <http://edlabs.ed.gov>.

Alternate Formats Upon request, this report is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

Disclosure of potential conflict of interest

The research team for this study was based at Regional Educational Laboratory West administered by WestEd. Neither the authors nor WestEd and its key staff have financial interests that could be affected by the findings of this study. No one on the 11-member Technical Working Group, convened annually by the research team to provide advice and guidance, has financial interests that could be affected by the study findings.¹

¹Contractors carrying out research and evaluation projects for IES frequently need to obtain expert advice and technical assistance from individuals and entities whose other professional work may not be entirely independent of or separable from the tasks they are carrying out for the IES contractor. Contractors endeavor not to put such individuals or entities in positions in which they could bias the analysis and reporting of results, and their potential conflicts of interest are disclosed.

Acknowledgments

This study and this report were made possible by a collaborative effort of many school principals, teachers, families, students, and researchers. The Regional Educational Laboratory (REL) West would like to thank the following colleagues who made the study possible.

First, we thank the teachers, families, students, and school and district administrative staff who made every step of program implementation and data collection possible. We recognize the burden associated with participating in a research study of this size and thank them for their time, commitment, diligence, and interest.

Young People's Press, the developer of Lessons in Character (along with Dr. B. David Brooks), was firmly committed to rigorously evaluating the program. We thank Robert Saielli and Patricia Pflum for supporting the implementation and evaluation and Michelle Gray for efficiently securing and delivering curriculum materials at a moment's notice to schools throughout the state. We acknowledge the unwavering commitment and support of Dr. B. David Brooks, who traveled throughout California to train and offer implementation support to participating teachers.

We thank Bianca Reyes of WestEd for her steadfast work in study coordination and data collection. We also thank Michal Clingman of WestEd for her persistence in following-up with school districts to secure archival student data. We also appreciate the large team of survey proctors and support staff who supported data collection: Jorge Cruz, Linsey Delaney, Quentin Graham, Zeta Heiter, Dianne Larson, Brian Lugo, Luisa Pena, Jeff Polik, Ray Polik, Anselmo Rocha, Veasna Tep, and Brandon Turner.

Finally, we appreciate the technical consultation and advice provided by Ellen Kisker of the Analytical and Technical Support group, and by members of the REL West Technical Working Group:.. Jamal Abedi, Lloyd Bond, Geoffrey Borman, Brian Flay, Tom Good, Corinne Herlihy, Joan Herman, Heather Hill, Roger Levine, Juliet Shaffer, and Jason Snipes.

Contents

| | |
|--|------------|
| Disclosure of potential conflict of interest | iii |
| Acknowledgments | iv |
| Executive summary..... | xi |
| Research questions | xi |
| Study context and design | xii |
| Key outcomes and measures | xiii |
| Intervention description | xiv |
| Analysis and results | xv |
| Implementation..... | xv |
| Estimated impacts | xv |
| Synopsis | xvi |
| 1. Introduction and study overview..... | 1 |
| Need for the study | 1 |
| Typical character education practices in elementary schools | 3 |
| Lessons in Character program..... | 4 |
| Previous research on Lessons in Character | 5 |
| Theory of action for Lessons in Character program | 6 |
| Research domains and study questions | 8 |
| Primary confirmatory research questions—student outcomes..... | 8 |
| Intermediate confirmatory research questions—school environment outcomes | 9 |
| Exploratory research questions | 9 |
| Organization of report | 10 |
| 2. Study design and methodology | 12 |
| Sampling and power estimates..... | 14 |
| Sample recruitment | 14 |
| Cohort 1 school recruitment | 15 |
| Cohort 2 school recruitment | 15 |
| Incentives to participate | 16 |
| Participant consent and consent follow-up..... | 16 |
| Random assignment | 17 |
| Sample selection and retention | 18 |
| School participation..... | 18 |
| Teacher participation..... | 19 |
| Student participation | 22 |
| Teacher and student migration and crossovers | 25 |
| Data collection | 26 |
| Primary and intermediate outcome data collection and measures | 28 |

| | |
|---|-----------|
| Implementation data collection and measures | 30 |
| Baseline data collection and measures | 31 |
| Response rates | 33 |
| Outcome domain measures | 36 |
| Sample characteristics | 36 |
| Data analysis methods..... | 43 |
| Sensitivity analyses | 45 |
| Multiple hypothesis testing procedures..... | 47 |
| Missing data | 48 |
| 3. Implementation of Lessons in Character intervention..... | 49 |
| Intended Lessons in Character program implementation..... | 49 |
| Training and coaching | 49 |
| Core and supplementary curricular material | 50 |
| Schoolwide Lessons in Character activities | 50 |
| Actual Lessons in Character program implementation | 50 |
| Training | 50 |
| Lessons in Character core and supplementary lesson implementation | 51 |
| Evidence of character development activities and artifacts | 55 |
| Informal character development programs and activities | 55 |
| Summary of implementation findings..... | 56 |
| 4. Impact results..... | 61 |
| Summary of impacts of Lessons in Character..... | 61 |
| Primary confirmatory impacts of Lessons in Character (student outcomes) | 61 |
| Primary research question 1 | 61 |
| Primary research question 2 | 62 |
| Primary research question 3 | 62 |
| Primary intermediate impacts of Lessons in Character (school environment) | 64 |
| Intermediate research question 4 | 64 |
| Intermediate research question 5 | 64 |
| Sensitivity analyses | 65 |
| 5. Exploratory results | 66 |
| Summary of exploratory impacts of Lessons in Character | 66 |
| Exploratory impacts of Lessons in Character (student outcomes)..... | 67 |
| Exploratory research question 6 | 67 |
| Exploratory research question 7 | 69 |
| Sensitivity analyses | 69 |
| 6. Summary of study findings and study limitations | 71 |
| Overview of findings | 71 |
| Study limitations | 72 |

| | |
|---|--------------|
| Appendix A. Statistical power estimates..... | A-1 |
| Design stage statistical power estimates | A-1 |
| Statistical power estimates based on final analytic sample..... | A-3 |
| Appendix B. Teacher and student recruitment and retention by grade level and data source..... | B-1 |
| Appendix C. Survey items used to measure confirmatory and intermediate outcomes..... | C-1 |
| Appendix D. Internal consistency reliability and intraclass correlations of confirmatory and intermediate outcomes measures..... | D-1 |
| Appendix E. School-level and teacher-level baseline characteristics for retained sample..... | E-1 |
| Appendix F. Covariates included in impact analysis models..... | F-1 |
| Student-level impact models | F-1 |
| Teacher-reported intermediate outcome models | F-1 |
| Appendix G. Sensitivity analyses—primary impacts on student-level outcomes | G-1 |
| Appendix H. Sensitivity analyses—primary impacts on teacher-level outcomes | H-1 |
| Appendix I. Sensitivity analyses—exploratory impacts on student-level outcomes | I-1 |
| Appendix J. Sensitivity analyses—exploratory impacts on teacher-level outcomes..... | J-1 |
| Appendix K. Teacher survey | K-1 |
| Appendix L. Student survey | L-1 |
| Appendix M. Teacher implementation log | M-1 |
| References..... | REF-1 |

Tables

| | |
|--|-----|
| Table 1. Impact analysis outcomes | xiv |
| Table 2.1. Study characteristics and data collection schedule for Lessons in Character impact evaluation | 13 |
| Table 2.2. Randomization strata characteristics | 18 |
| Table 2.3. Teacher participation and mobility rates of participating teachers by experimental group | 21 |
| Table 2.4. Student participation rates by sample inclusion criterion and experimental group..... | 25 |
| Table 2.5. Students who migrated across schools by crossover status | 26 |
| Table 2.6. Measurement matrix of primary and intermediate outcome variables | 27 |
| Table 2.7. Data collection schedule | 27 |
| Table 2.8. Data collected by source, school grade, and year | 28 |

| | |
|--|-----|
| Table 2.9. Response rates by data source and experimental group | 35 |
| Table 2.10. School-level baseline characteristics of study schools (randomized sample) | 37 |
| Table 2.11. Teacher baseline demographic characteristics by experimental group (teacher participants in baseline year) | 38 |
| Table 2.12. Teacher baseline measures by experimental group (teacher participants) | 39 |
| Table 2.13. Student demographic characteristics by experimental group | 40 |
| Table 2.14. Student baseline measures by experimental group | 42 |
| Table 3.1. Lessons in Character training attendance | 51 |
| Table 3.2. Lessons implemented by grade 2-5 teachers in intervention group schools..... | 52 |
| Table 3.3. Lessons implemented by grade 4-5 teachers in intervention group schools..... | 53 |
| Table 3.4. Teacher-reported implementation fidelity in intervention group schools (grade 2-5 teachers) | 54 |
| Table 3.5. Teacher-reported implementation fidelity in intervention group schools (grade 4-5 teachers) | 55 |
| Table 3.6. Observed artifacts indicative of social and character education programs in intervention and control group schools (proctor observation checklist)..... | 57 |
| Table 3.7. Social development activities implemented in study classrooms and schools (grades 2-5, past six months) | 59 |
| Table 3.8. Social development activities implemented in study classrooms and schools (grades 4-5, past six months) | 60 |
| Table 4.1. Impact analysis of primary outcome variables, primary analytic sample second-year impacts | 63 |
| Table 4.2. Impact analysis of intermediate outcome variables, primary intermediate analytic sample second-year impacts | 65 |
| Table 5.1. Impact analysis of exploratory outcome variables, exploratory analytic sample first-year impacts | 68 |
| Table 5.2. Impact analysis of exploratory outcome variables, exploratory analytic sample first-year impacts (grades 2-5) | 69 |
| Table A1. Parameters used to estimate statistical power in planning phase and based on achieved sample | A-2 |
| Table A2. Minimum detectable effect size estimates for student and teacher outcome measures..... | A-3 |
| Table D1. Cronbach alpha reliability coefficients and school intraclass correlations of student measures addressing key research questions..... | D-1 |
| Table D2. Cronbach alpha reliability coefficients and intraclass correlations of teacher measures addressing intermediate research questions | D-1 |
| Table E1. Teacher baseline demographic characteristics by experimental group | E-1 |
| Table F1. Covariates included in academic performance impact analysis models..... | F-2 |
| Table F2. Covariates included in social competence impact analysis models | F-3 |
| Table F3. Covariates included in problem behavior impact analysis models..... | F-4 |
| Table F4. Covariates included in school environment impact analysis models (students) | F-5 |
| Table F5. Covariates included in school environment impact analysis models (teachers) | F-6 |
| Table G1. Estimated impact on English language arts test scores across different models, samples, and methods used to handle missing data (primary analytic sample)..... | G-1 |
| Table G2. Estimated impact on academic competence across different models, samples, and methods used to handle missing data (primary analytic sample)..... | G-2 |

| | |
|---|------|
| Table G3. Estimated impact on Social Skills Rating System across different models, samples, and methods used to handle missing data (Primary Analytic Sample) | G-3 |
| Table G4. Estimated impact on altruism across different models, samples, and methods used to handle missing data (primary analytic sample) | G-4 |
| Table G5. Estimated impact on empathy across different models, samples, and methods used to handle missing data (primary analytic sample) | G-5 |
| Table G6. Estimated impact on externalizing behavior across different models, samples, and methods used to handle missing data (primary analytic sample) | G-6 |
| Table G7. Estimated impact on aggression across different models, samples, and methods used to handle missing data (primary analytic sample) | G-7 |
| Table G8. Estimated impact on delinquent behavior across different models, samples, and methods used to handle missing data (primary analytic sample) | G-8 |
| Table G9. Estimated impact on school expectations across different models, samples, and methods used to handle missing data (primary analytic sample) | G-9 |
| Table G10. Estimated impact on student feelings of belonging across different models, samples, and methods used to handle missing data (primary analytic sample) | G-10 |
| Table H1. Estimated impact on teacher reports of school expectations across different models, samples, and methods used to handle missing data (primary analytic sample – grades 2-5) | H-1 |
| Table H2. Estimated impact on teacher reports of student feelings of belonging across different models, samples, and methods used to handle missing data (primary analytic sample – grades 2-5) | H-1 |
| Table H3. Estimated impact on teacher reports of school expectations across different models, samples, and methods used to handle missing data (primary analytic sample – grades 4-5) | H-2 |
| Table H4. Estimated impact on teacher reports of student feelings of belonging across different models, samples, and methods used to handle missing data (primary analytic sample – grades 4-5) | H-2 |
| Table I1. Estimated impact on English language arts test scores across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-1 |
| Table I2. Estimated impact on academic competence across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-2 |
| Table I3. Estimated impact on Social Skills Rating System across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-3 |
| Table I4. Estimated impact on altruism across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-4 |
| Table I5. Estimated impact on empathy across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-5 |
| Table I6. Estimated impact on externalizing behavior across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-6 |
| Table I7. Estimated impact on aggression across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-7 |
| Table I8. Estimated impact on delinquent behavior across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-8 |
| Table I9. Estimated impact on school expectations across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-9 |

| | |
|--|------|
| Table I10. Estimated impact on student feelings of belonging across different models, samples, and methods used to handle missing data (exploratory analytic sample) | I-10 |
| Table J1. Estimated impact on teacher reports of school expectations across different models, samples, and methods used to handle missing data (exploratory analytic sample – grades 2-5) | J-1 |
| Table J2. Estimated impact on teacher reports of student feelings of belonging across different models, samples, and methods used to handle missing data (exploratory analytic sample – grades 2-5)..... | J-1 |
| Table J3. Estimated impact on teacher reports of school expectations across different models, samples, and methods used to handle missing data (exploratory analytic sample – grades 4-5) | J-2 |
| Table J4. Estimated impact on teacher reports of student feelings of belonging across different models, samples, and methods used to handle missing data (exploratory analytic sample – grades 4-5)..... | J-2 |

Figures

| | |
|--|-----|
| Figure 1.1. Theory of action for Lessons in Character program..... | 7 |
| Figure 2.1. Flow diagram of nonparticipating and participating teachers into and out of study schools..... | 20 |
| Figure 2.2. Flow diagram detailing student analytic samples..... | 24 |
| Figure B1. Flow diagram of recruitment and retention of schools and study participants | B-1 |
| Figure B2. Flow diagram of teacher recruitment and retention..... | B-2 |
| Figure B3. Flow diagram of student recruitment and retention..... | B-3 |

Boxes

| | |
|---|-----|
| Box C1. Items used to assess primary and exploratory student outcomes | C-1 |
| Box C2. Items used to assess intermediate and exploratory school environment outcomes | C-2 |

Executive summary

Character education has become one of the fastest growing reform movements in K–12 education today, partially in response to unacceptable levels of student misbehavior and inadequate endorsement of good character values (Williams 2000). Between 1993 and 2009, 36 states passed laws mandating or recommending some aspect of character education in schools. Character education programs also have strong support from parents, teachers, and school administrators (Character Education Partnership 2010; Glanzer and Milzen 2006). But despite the widespread popularity of such programs, relatively few randomized controlled trials have examined the impact of these programs on the character traits, behavior, and academic outcomes of students (What Works Clearinghouse 2007).²

This study examines the impact of the Lessons in Character (LIC) program—an English language arts–based character education program—on student academic achievement, social competence, and problem behaviors and, secondarily, on the school environment. The program consists of literature-based supplementary material aligned with California English language arts standards and designed to integrate easily into the current English language arts curricula. The LIC curriculum is designed to be easy to implement in the classroom and to involve minimal teacher training, which distinguishes the program from other character education programs.

Research questions

The following confirmatory research questions guide this study:

1. Do students in grades 4 and 5 who attend schools in the LIC intervention group exhibit higher scores on measures of academic achievement after two academic years of potential LIC exposure than their counterparts who attend schools in the control group?
2. Do students in grades 4 and 5 who attend schools in the LIC intervention group exhibit higher scores on measures of social competence after two academic years of potential LIC exposure than their counterparts who attend schools in the control group?
3. Do students in grades 4 and 5 who attend schools in the LIC intervention group exhibit fewer problem behaviors after two academic years of potential LIC exposure than their counterparts who attend schools in the control group?

To examine part of the action model linking the LIC program to student outcomes, the study also examines whether participation in the program is associated with enhancements to two areas of the school environment. Specifically, the following two confirmatory research questions involving intermediate outcomes are examined:

4. Do teachers and students in the LIC intervention group report greater levels of student belongingness after two years of program implementation than their counterparts in the control group?

² A recent exception is the evaluation of seven social and character development programs conducted by the Social and Character Development Research Consortium (2010).

5. Do teachers and students in the LIC intervention group report greater levels of school expectations consistent with character development after two years of program implementation than their counterparts in the control group?

The confirmatory research questions focus on the impacts of the LIC program after a two-year implementation period.

Exploratory analyses were conducted to examine potential impacts on the same outcome measures after one year of program implementation. These analyses were exploratory in the sense that they were used to obtain preliminary estimates of program impacts that could not be rigorously tested due to statistical power considerations. The first-year results were not used to make inferences about the overall effectiveness of the intervention but to provide additional descriptive information.

Study context and design

Designed as an experimental trial, the study took place from spring 2007 to spring 2010 in 50 California elementary schools with teachers of grades 2–5. This implementation period corresponded with an economic recession in California (and the nation as a whole) and substantial school spending reductions and teacher layoffs due to state budget shortfalls in California.

Participation in the study was voluntary. The intervention involved a one-day training course for teachers in year 1; teachers who participated in the course were expected to incorporate 19–25 supplementary lessons into their classroom instruction each year for two academic years. No LIC professional development activities or coaching occurred in year 2 of implementation.³ Two cohorts of elementary schools participated in the intervention. Cohort 1 implementation took place in 2007/08 and 2008/09, with teacher professional development and coaching occurring in late summer and early fall of year 1 (2007). Cohort 2 schools were recruited in spring 2008, with teacher professional development occurring in summer/fall 2008/09 and 2009/10. Teachers in control group schools, meanwhile, continued with their regular professional development activities and instructional practices. Both cohorts were pooled in the data analyses.

The study was conducted in 50 public elementary schools in California—34 in the Los Angeles and San Diego Metropolitan Statistical Areas, 13 in northern California, and 3 in central California. Half of the recruited schools were randomly assigned to an intervention group that had the opportunity to implement LIC in their 2nd–5th grade classrooms for two consecutive years, and half were assigned to the control group. The primary impact analyses included 4,683 students who were in grade 4 or 5 in year 2, and the exploratory analyses included 5,674 students who were in grade 4 or 5 in year 1. Students in the primary analytic sample were present and obtained parent consent to participate by the beginning of year 1 and were in participating classrooms when spring year 2 data were collected. Those in the exploratory analytic sample obtained parent consent by the beginning of year 1 and were in participating classrooms when spring year 1 data were collected. Students who were enrolled in non-participating teachers' classrooms—460 students in year 1 and 300 students in year 2—were not included in the analytic samples.

³ Teachers who joined the study in year 2 did not receive any training on the use of the curriculum.

Thirty schools were in a large city or on the urban fringe of a large city, six in a mid-size city (population of 25,000–250,000), and fourteen in a small town or rural area. Each school served an average of 354 students. Nine schools had fewer than 150 students, and two had more than 600 students. Approximately 59 percent of the students were eligible for free or reduced-price lunch, 49 percent were classified as Hispanic, 32 percent were classified as non-Hispanic White, and 31 percent were classified as English language learner students (California Department of Education 2009).

With an average of 96 grade 4 or 5 students in each school, the sample size is sufficient for detecting program impacts on student outcomes of 0.27 standard deviations for primary academic outcomes and 0.20 standard deviations for primary social competence and behavioral outcomes. Program impacts of this magnitude are equivalent to about two-thirds of a year of growth experienced by grade 4 students on norm-referenced standardized tests in reading (Hill et al. 2008).

Key outcomes and measures

Table 1 lists the study's key outcome variables—student academic achievement, social competence, problem behaviors, and the school environment—and their measures. These outcomes were assessed with standardized achievement tests, teacher reports on the Social Skills Rating System (Gresham and Elliott 1990), surveys of grade 4 and 5 students, and teacher surveys.

- *State English language arts assessments.* Student achievement data from state-mandated standardized assessments of English language arts (the California Standards Tests) were collected for the years before and during program implementation. Criterion-referenced to state standards, the California Standards Tests in English language arts are administered to students in grades 2–11.
- *Social Skills Rating System teacher reports.* Gresham and Elliott's (1990) Social Skills Rating System was used to assess student social skills, problem behaviors, and academic competence.
- *Student surveys.* A 35-minute survey assessing behaviors, attitudes, and values consistent with the goals of character education was administered to all grade 4 and 5 students in the fall and spring of year 1 and the spring of year 2. Using items and subscales from validated instruments, the survey assessed student altruism (Characterplus 2002), aggression (Orpinas and Frankowski 2001), delinquent behavior (Kisker et al. 2004), and empathy (Funk et al. 2003), as well as school belonging and expectations (Characterplus 2002).
- *Teacher surveys.* All teachers were surveyed in the spring of year 0 before schools were randomly assigned to the intervention or control group. Teachers were surveyed again in the spring of years 1 and 2. The year 0 survey assessed teacher background and school climate—teacher education, professional development experiences in language arts and character education, activities related to social and character development that teachers implemented in their classrooms, and school and classroom climate (students' feelings of belonging and school expectations). The year 1 and year 2 surveys also contained

questions that asked about other professional development that teachers participated in during program implementation and questions that assessed teachers' communication and collaboration.

Table 1. Impact analysis outcomes

| Outcome | Measure |
|--|-----------------------------|
| <i>Academic achievement</i> | |
| English language arts standardized tests | State assessment |
| Academic competence | Teacher report |
| <i>Social competence</i> | |
| Social Skills Rating System total score | Teacher report |
| Altruism | Student survey |
| Empathy | Student survey |
| <i>Problem behaviors</i> | |
| Externalization of problems | Teacher report |
| Aggression | Student survey |
| Delinquent behavior | Student survey |
| <i>School and classroom climate</i> | |
| Students' feeling of belonging | Teacher and student surveys |
| School expectations | Teacher and student surveys |

Intervention description

Developed by Dr. B. David Brooks, with support from Young People's Press, the LIC program is delivered by classroom teachers, with implementation support from Dr. Brooks. Through multicultural literature (lap books) and audiocassettes, it teaches character and integrates the language of character into the English language arts curricula. The LIC program is designed for grades K–9 and aligned with California education standards. Teachers in schools randomly assigned to the intervention condition participated in a one-day training session, and, in the fall semester, received approximately two hours of coaching support for delivering the curricular material.

The curricular material comprises two components, both designed to reinforce good character and support language arts learning standards: the core curriculum, *Lessons in Character*, and the supplementary materials, *Daily Oral Language with Character* and *Writing with Character*. The LIC core curriculum explicitly integrates the language of character into the curriculum with a focus on teaching for understanding that emphasizes civility, respect, responsibility, trustworthiness, fairness, caring, loyalty, and self-control. It is a literature-based language arts program that relies on multicultural literature, enrichment activities, cross-curricular activities, read-aloud books, and questioning, all of which are intended to help instill habits of good character. The core curriculum consists of 25 20-minute lessons; study teachers are asked to implement at least 19 lessons during the academic year. It begins with a decision-making model—Stop, Think, Act, Review (STAR)—that is used throughout the year as a classroom management tool and reinforcement of the program lessons.

According to the developer, delivery of the LIC lesson materials does not require extensive teacher preparation or training. Each teacher is provided with a teacher management guide that

provides concise lesson plans for each of the 25 lessons and examples of discussion questions and classroom activities.

The supplementary materials are optional for teachers. The Daily Oral Language with Character materials consist of sentence correction activities followed by short writing assignments that emphasize decision-making, goal setting, civic responsibility, and other character education components. Designed for daily use, Daily Oral Language with Character augments the regular language arts program with lessons that take no longer than five minutes. The Writing with Character materials (for grades 3–8) consist of 36 weekly 20-minute writing assignments that focus on the mechanics of writing (e.g., grammar, spelling, punctuation) as well as on character education.

Although the LIC program is a comprehensive schoolwide character education program, it is the program's integration into the curriculum that is intended to secure teacher support. Teachers become program experts through daily teaching, and this acquired expertise helps build support for comprehensive schoolwide character education policies and practices.

Analysis and results

To estimate program impacts, outcomes for students and classrooms in intervention group schools were compared with those for students and classrooms in control group schools.

Multilevel regression models were used to estimate the effects of the LIC program and to account for data clustering by school and classroom (Goldstein 1987; Raudenbush and Bryk 2002; Murray 1998). The impact analyses controlled for baseline (pretest) measures of outcome variables and other student- and school-level covariates.

Implementation

The LIC core curriculum consists of 25 lessons (24 core principal lessons and one STAR lesson), of which intervention group teachers were asked to implement 19 during the academic year. Teachers implemented an average of 12.40 LIC lessons in year 1 and 9.56 LIC lessons in year 2. In year 1, 16 percent of teachers reported that they did not implement any lessons, while 40 percent implemented 16 or more lessons. Teachers' reported use of the curricular materials declined after year 1: in year 2, 28 percent did not implement any LIC lessons, and 28 percent implemented 16 or more lessons. The supplementary curricular materials were used less frequently than the LIC core materials: in year 1, two-thirds of teachers used at least some of the Daily Oral Language with Character or Writing with Character materials in their classrooms, and in year 2, about half of teachers reported using these materials.

Estimated impacts

The primary confirmatory impact analyses indicated that grade 4 and 5 students who attended schools in the LIC program intervention group did not exhibit higher scores on measures of academic achievement and social competence, or lower scores on measures of problem behaviors after two academic years of potential LIC exposure than grade 4 and 5 students who attended schools in the control group. Moreover, the intermediate impact analyses indicated that there were no statistically significant LIC program impacts on the school environment measures of school expectations and student feelings of belonging. Exploratory analyses suggested that

there were no statistically significant LIC program impacts on grade 4 or 5 student outcomes or on measured school environment outcomes after the first year of program implementation.

Synopsis

The impact analyses did not find that LIC improved student academic achievement, social competence, or problem behaviors on any of the analyzed measures. Nor were statistically significant impacts on the school environment measures of expectations and student belonging detected. Moreover, analyses of teacher reports of program implementation indicated that 30 percent of teachers reported that, in year 1, they implemented the number of core lessons recommended by the developer and that, in year 2, 23 percent reported having done so. The failure to detect program impacts on student and school environment outcomes could be due to the weak implementation of the program, an intervention design that simply does not work in bringing about changes in student and school environmental outcomes, or methodology limitations of the study.

An important limitation of the study was that retention rates differed between teachers in intervention and control schools. Although participation rates in intervention and control schools did not differ and there were no statistically significant differences in the baseline characteristics of teachers or students in the intervention and control schools, it is possible that there were important unobserved differences between intervention and control samples that could have resulted in biases in estimated program impacts.

1. Introduction and study overview

This study examines the impact of the Lessons in Character (LIC) program—an English language arts-based character education program—on student academic achievement, social competence, problem behaviors, and school climate. Since 1995, the LIC program has been implemented in more than 15,000 schools and in every state except Alaska. The core of the LIC program is literature-based curricular material designed to be integrated easily into an existing English language arts curriculum. The lessons are aligned with California English language arts standards. The infusion of LIC lessons into existing English language arts curricula, and the resulting ease of implementation, distinguishes the program from other character education programs.

Need for the study

Character education programs have become increasingly popular in K–12 education, partially in response to ostensibly high levels of student misbehavior and concerns about low levels of endorsement of values consistent with good character (Williams 2000). Between 1993 and 2009, 36 states passed laws mandating or recommending some aspect of character education in schools, and character education programs have high levels of support from parents, teachers, and school administrators (Character Education Partnership 2010; Glanzer and Milzen 2006). The No Child Left Behind Act of 2001 enacted the Partnerships in Character Education Program to support the design and implementation of instruction aimed at promoting positive character development and improving the school environment. Over 2001–06, Congress appropriated approximately \$25 million a year in grants to states and school districts to design and implement character education programs (U.S. Department of Education 2010). Although U.S. Department of Education appropriation levels for the program have since declined, interest in character education among K–12 educators is high, a wide variety of character education curricula and professional development resources are available, and implementation of character education programs in schools is pervasive (Berkowitz and Bier 2007).

The primary rationale for character education is to promote the ethical, social, and personal integrity of students. Proponents of character education argue that the nation benefits when its citizens subscribe to the ideals of respect for others, fairness, justice, honesty, responsibility, and civic participation (Ryan 2003). Character education programs are also promoted as a partial solution to student misbehavior at school and the effect of such misbehavior on student learning (Benninga et al. 2003; Ryan 2003). Correlational evidence drawn from years of research has shown that adolescent substance use, violence, crime, and antisocial behavior are associated with low levels of academic achievement and other school-related factors, including reduced attention spans, lower investment in homework, more negative attitudes toward school, less motivation, and increased absenteeism (Hanson, Austin, and Bayha 2004). Such factors may also adversely affect academic performance by influencing teaching and learning processes in the classroom. For example, Lochman et al. (1987) found that students who were disruptive and aggressive in the classroom divert teachers' attention and reduce instruction time, thereby potentially reducing their classmates' academic achievement (Bowen and Bowen 1999).

Despite the widespread popularity of character education programs, relatively few prospective randomized controlled trials have been conducted to examine the impact of these programs on

the character traits, behavior, and academic outcomes of students (What Works Clearinghouse 2007). In 2006, the What Works Clearinghouse identified 93 studies evaluating 41 character education program. Of those, seven studies used an experimental design that met evidence standards, one relied on an experimental design that had confounding problems, and nine utilized a quasi-experimental with a well-matched comparison group (What Works Clearinghouse 2007).

Notably, the Social and Character Development (SACD) Research Consortium (2010) recently completed a rigorous, multisite, experimental study that assessed the effectiveness of seven school-based social and character development programs implemented in elementary schools. The evaluated programs used character education, socioemotional learning, violence prevention, and behavior management strategies to improve student behavior. The programs evaluated include the Academic and Behavioral Competencies Program; the Competence Support Program; Love in a Big World; Positive Action; Promoting Alternative Thinking Strategies (PATHs); Reading, Writing, Respect, and Resolution (4Rs); and Second Step. Each program's planned activities were integrated and logically organized based on a theory of action, were implemented in schools by teachers and other school personnel (school-based), and were directed at all students in participating schools (universal). The SACD evaluation found no evidence that any of the seven programs improved students' social and character development—either collectively across all programs, individually by program, or for specific subgroups of students.

A critical finding of the SACD evaluation was the high level of implementation of activities designed to promote students' social and character development in control schools. As described above, many states require or promote social and character development activities in schools (Character Education Partnership 2010; Glanzer and Milzen 2006), and schools participating in impact evaluations of social and character development programs are likely to be highly interested in implementing such activities. The SACD Research Consortium reported that between 86 percent and 90 percent of teachers in control schools reported implementing activities to promote at least one of six social and character development goals.⁴ Although teachers in intervention group schools reported higher use of such activities in their classrooms, differences in practices across intervention and control schools may not have been great enough to generate detectable impacts on student outcomes.

Other evaluation studies of several of the programs examined by the SACD Research Consortium have reported positive impacts on students' social competence, behavior, and academic achievement. For example, another randomized controlled trial that investigated the effectiveness of the Positive Action program found evidence that the program reduced students' substance use, violence, voluntary sexual activity, absenteeism, suspension, and grade retention; and increased students' scores on standardized tests in English language arts and mathematics (Beets et al. 2009; Snyder et al. 2010). A randomized efficacy trial of the PATHs program—which focuses on enhancing student self-control, emotional awareness, peer-related social skills, and social problem solving (Flay, Berkowitz, and Bier 2009)—found beneficial impacts in the areas of student aggression and social competence (Conduct Problems Prevention Research Group 2010).

⁴ These six social and character development goals are: (1) violence prevention and peace promotion, (2) social and emotional development, (3) character education, (4) tolerance and diversity, (5) risk prevention and health promotion, and (6) civic responsibility and community service.

In addition, Jones and her colleagues' (2011) examination of the impacts of the 4Rs (Reading, Writing, Respect, and Resolution) Program suggests that this intervention may be effective in improving student outcomes in areas and for subgroups that were not the focus of the SACD evaluation. Focusing on a set of outcomes (assessed in year 2) that were central to the program-specific theory of change, Jones et al. found that students in schools that were randomly assigned to the intervention condition exhibited improvements in child reports of depression and teacher reports of attention/hyperactivity symptoms, aggression, and social competence relative to their counterparts in control group schools. Moreover, the 4Rs Program had statistically significant, positive impacts on mathematics and reading achievement among students with high levels of aggression and conduct problems at baseline.

Like the LIC program, 4Rs uses a literature-based curriculum to integrate social and character development into the English language arts curriculum. The curricular component of 4Rs is comprised of seven units and 21 lessons. Each unit is organized around a specific children's book that is used to enhance student understanding and skills in listening, handling anger, cooperation, and negotiation (Aber et al. 2011). Unlike the LIC program, 4Rs provides 25 hours of intensive professional development followed by ongoing classroom coaching (12 times per year) to teachers. A critical component of 4Rs logic model is teachers as mechanisms of change. The program aims to inspire teachers to employ the ideas and practices embodied in the program in their own lives as well as in their classrooms (Aber et al. 2011; Flay, Berkowitz, and Bier 2009). The alignment of teachers' values, beliefs, and perceptions with the principles of the program are thought to be critical for effective implementation (Flay, Berkowitz, and Bier 2009).

Typical character education practices in elementary schools

Character education programs typically employ strategies that have much in common with other types of programs, including socioemotional learning, life skills, violence prevention, and substance use prevention programs. However, although "character education" has been defined broadly to encompass any program designed to promote positive youth development (Berkowitz and Bier 2007; Battistich 2005), what distinguishes character education programs from other programs that focus on student social skills and behavior is an emphasis on instilling and promoting a range of core values, such as trustworthiness, fairness, and responsibility (What Works Clearinghouse 2007). In describing character education, this study focuses on programs that explicitly focus on directly teaching core values.

A diverse set of character education initiatives and programs has been developed for K–12 education. These programs vary in how much they focus on development of social and emotional skills, are integrated into the academic curriculum, are integrated into the full spectrum of school activities and school stakeholders (such as students, staff, and family), and use direct and interactive teaching strategies (Berkowitz and Bier 2007; Benninga et al. 2003). Little information has been systematically collected on the scope of use of specific character education programs or types of programs. The most commonly implemented character education program is likely Character Counts!, which is designed as a framework for implementation of character development strategies rather than a packaged curriculum. Character Counts! is based on a set of values referred to as the "Six Pillars of Character" (trustworthiness, respect, responsibility, fairness, caring, and citizenship; see Aspen Declaration on Character Education 1992). It focuses on direct instruction about core values, using classroom-based and schoolwide activities such as

implementing literature-based character lessons in classrooms, including character messages in morning announcements, organizing service projects, and publicly recognizing students who display behaviors consistent with values included in the Six Pillars. Character Counts! is designed to be flexible and easily integrated into teachers' existing curriculum and classroom management practices. To date, no experimental or quasi-experimental, comparison- group studies have been conducted to examine the effectiveness of Character Counts! on student outcomes.

An example of a packaged character education program is the Heartwood Institutes' An Ethics Curriculum for Children—a read aloud, literature-based curriculum designed to teach children about seven values: courage, loyalty, justice, respect, hope, honesty, and love (Lemming 2002). Each of the 14 lessons consist of an introduction of the concept, a story read to the students by the teacher that illustrates the concept, discussion questions that help explain and apply the concept, and activities to reinforce the concept. Like the LIC program, the goal of the curriculum is to promote student understanding, endorsement, and behavioral enactment of the seven ethical values. A What Works Clearinghouse review of evidence regarding effectiveness found that An Ethics Curriculum for Children had no discernable effects on student behavior, attitudes, or values (What Works Clearinghouse 2007).

The Love in a Big World (LBW) program (one of the programs included in the SACD evaluation discussed above) also uses story reading as one strategy to make character traits more tangible to children. Based on research conducted by the Social Development Research group at the University of Washington and on youth asset development research conducted by the Search Institute, the LBW program aims to create a more caring climate within the classroom and school, provide instruction to students on character traits, and provide opportunities for bonding (Flay, Berkowitz, and Bier 2009). Program elements include classroom lesson plans, a peer recognition program, and school-wide strategies for teaching character traits. Daily 10- to 15- minute classroom lessons use story reading, writing exercises, discussion, songs, and other activities to increase knowledge about and endorsement of character traits. Teachers are instructed to model and reinforce social skills throughout the day and to utilize behavioral management strategies to reward students' demonstration of character traits. Schoolwide strategies include weekly announcements, assemblies, service projects, parent newsletters, and visual artifacts that illustrate character traits. As a complete package, LBW program components are designed to increase student knowledge about character traits, improve social and emotional competence, increase student support for character development, and enhance student self-efficacy to practice character development. The SACD evaluation found no evidence that LBW impacted school climate or students' social and character development. No other evaluations of LBW have been published.

Lessons in Character program

The LIC program was developed by Dr. B. David Brooks, director of the Center for Implementing Character Education, with support from Young People's Press. Delivered by classroom teachers, with implementation support from Dr. Brooks, it is a literature-based language arts program that uses a collection of multicultural literature (lap books or large-format mini-anthologies) and audiocassettes to teach lessons of character. Designed for grades K–9, it

focuses on integrating the language of character into the English language arts curriculum. The curriculum is aligned with California education standards.

The curricular material comprises two components: the core curriculum, LIC, and the support materials, Daily Oral Language with Character and Writing with Character. All components are designed to reinforce the understanding of character traits and the enactment of related behaviors and to support language arts learning standards. The core curriculum and the support materials focus explicitly on trustworthiness, respect, responsibility, fairness, caring, and citizenship (the Six Pillars of Character).

The most central and critical component of the program is the LIC core curriculum. It is the component of the program that focuses on teaching for understanding and explicitly integrates the language of character into the curriculum. The pedagogical approach relies on multicultural literature, enrichment activities, cross-curricular activities, read-aloud books, and questioning strategies to help instill habits of reflection and consideration of the needs of others, which are associated with enacting traits of good character (Jolliffe and Farrington 2004; Warden and Mackinnon 2003). A “theme of the month” approach is used to engage students in targeted activities. LIC consists of 25 twenty-minute lessons designed to be delivered weekly during the course of the academic year. The program begins with an introduction to the decision-making model that is used throughout the year as a classroom management tool and reinforcement of the LIC lessons.

The Daily Oral Language with Character materials consist of sentence correction activities followed by short writing assignments that emphasize decision-making, goal setting, civic responsibility, and several other character education components. The materials are designed to be used daily to augment the regularly taught language arts program. Each lesson is designed to take no more than 5 minutes. The Writing with Character materials (for grades 3–8) consist of 36 weekly 20-minute writing assignments that allow students to practice composition and writing skills. The lessons focus on both the mechanics of writing (e.g., grammar, spelling, punctuation) and themes of character education.

School principals receive the *LIC School Administrator Implementation Guide*, which provides a “theme of the month” approach to presenting displays and other types of communication about character traits throughout the school. Posters, letters to families, and daily announcements are organized by character theme. Four how-to booklets offer strategies for implementing character education through guidance and community, family, and school activities. Although principals are encouraged to implement schoolwide aspects of the LIC program and are provided with information on how to implement the program, the training and support provided by the developer places more emphasis on teacher activities. The programmatic philosophy of LIC is that teachers become sufficiently engaged and proficient in implementing program principles with rigor and fidelity through infusing character education into the weekly curriculum. The developer of the curriculum believes that building teachers’ expertise in LIC through teaching helps build support for schoolwide, comprehensive character education policies and practices.

Previous research on Lessons in Character

To date, one randomized controlled trial has been conducted to investigate the short-term effectiveness of the LIC program. Based on a sample of 372 grade 4 students in Louisiana and

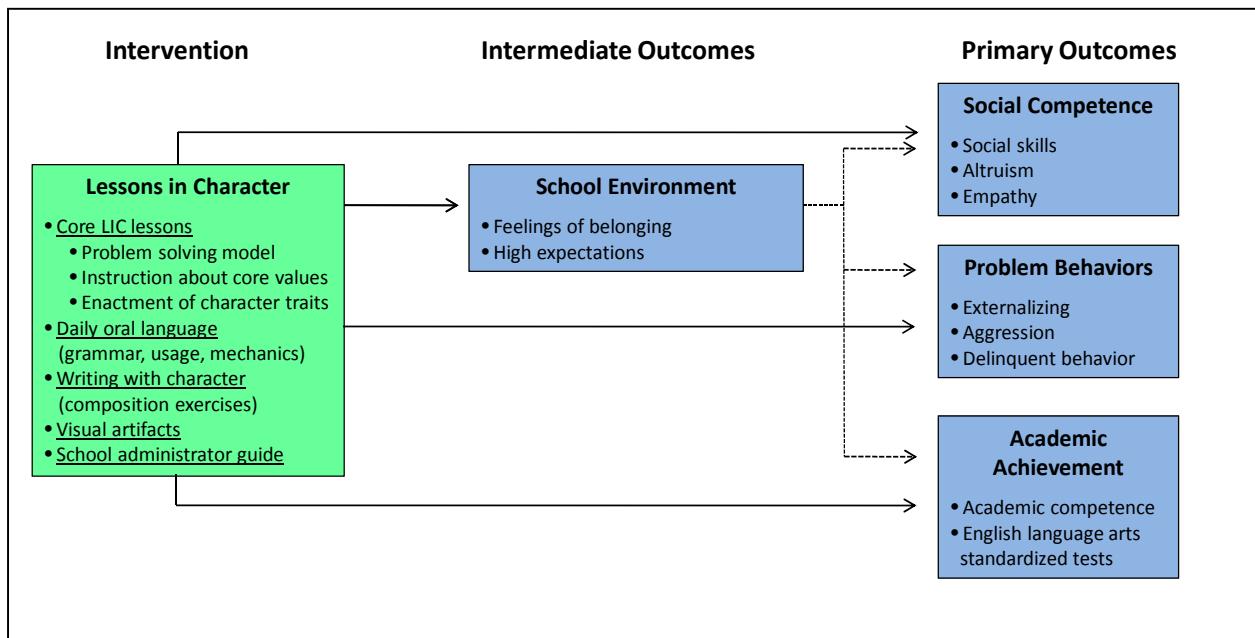
Florida, Dietsch, Bayha, and Zheng (2005) compared outcomes for 11 classrooms randomly assigned to the LIC curriculum for one semester with 10 randomly assigned control group classrooms. The study found statistically significant results favoring the intervention group on attendance, reading grades, and mathematics grades, with effect sizes ranging from 0.31 to 0.48 standard deviations. Item-level analyses of the student surveys found more favorable results for the intervention group on character-related knowledge, attitudes, and values, but only a few differences were statistically significant. Similarly, item-level analyses of teacher perceptions of student knowledge and behavior generally favored the intervention group, but again only a few differences were statistically significant. No statistically significant results were found that favored the control group.

Theory of action for Lessons in Character program

Figure 1.1 depicts the theoretical framework that guides the evaluation. The LIC program is designed to enhance student problem solving skills using an explicit decision-making process; promote student understanding, endorsement, and behavioral enactment of core values; and boost language, grammar, mechanics (punctuation, spelling), and composition skills through two supplementary English language arts workbooks. The language arts workbooks are designed to further reinforce core values and social competence by incorporating character education themes.

The LIC core curriculum begins with a lesson focusing on the S.T.A.R. (Stop, Think, Act, and Review) four-step decision-making model. The S.T.A.R. process emphasizes taking the appropriate amount of time to reflect about consequences of behavior before acting, making a mental list of options and consequences of alternative behavioral choices for oneself and others, and acting in accordance with one's examination of alternatives. After taking action, the process ends with student reflection about the consequences of the behavior, with an explicit focus on whether the behavior resulted in the student getting closer or farther from her/his goals and the consequences of the behavior on others. S.T.A.R. is intended to provide students with a tool to deliberately consider core values and the benefits of acting in accordance with core values as they engage in behavior.

Figure 1.1. Theory of action for Lessons in Character program



The remaining LIC core curriculum lessons focus on six core values: trustworthiness, respect, responsibility, fairness, caring, and citizenship. For each core value, four 20-minute lessons are included, always in this order: defining the theme and discussing why student acquisition of the character trait is desirable, focusing on applying the character trait in school and how doing so will benefit the student, focusing on commitment to the character trait in one's personal life and how such a commitment can help the student to be a better person, and focusing on applying the character trait in home and community settings and how this benefits the student as a family and community member and as a future worker (Brooks 1996). The classroom strategies used for teaching about core values include reading stories, classroom discussion activities, role-playing, writing exercises, class projects, and other classroom activities.

The stories that are used in LIC are infused with situations illustrating core values, with a focus on how the behaviors each individual in the story chooses influence outcomes. The teacher manual provides strategies and prompts for students using the Predictions in Reading Model (Hoskisson and Garrison 1998). For example, prompts during reading the stories might ask about the choices available to the individual in the story and the consequences of choosing specific behaviors. All of the LIC stories focus on one or two specific core values and provide opportunities for discussion about how the actions of the individuals in the story reflect those values. The stories provide students an opportunity to predict outcomes based on choosing behaviors linked to core values.

The theory of action holds that by providing instruction about what it means to enact core values, why doing so is desirable, and training students to use a decision-making procedure to put into practice what has been learned—students develop the understanding and skills they need to engage in pro-social behavior and avoid negative behavior.

To reinforce what is learned in the lessons and to create a school climate supportive of behaving in accordance with character traits, visual artifacts (posters) illustrating character themes and S.T.A.R are displayed throughout the school. Other schoolwide strategies include daily

announcements, letters to families, and guidance on integration of character education with the school discipline policies.

The components of the intervention are postulated to influence students' social competence, problem behaviors, and academic achievement both directly and indirectly via the school environment. The core LIC components of training students in a problem solving model, and training and promotion of student understanding, endorsement, and enactment of core values such as responsibility and kindness are postulated to improve the school environment, as manifest in increased feelings of belonging among students, and perceptions on the part of teachers and students that there are clearly articulated and high expectations for student behavior within the school. By promoting core values and enhancing problem solving skills, the intervention is also posited to increase social competence and reduce antisocial behavior. The language arts components of the intervention are expected to have direct effects on English language arts performance. Moreover, the LIC program may also enhance student learning and achievement through its effects on the skills and habits necessary for academic achievement—traits and behaviors such as responsibility, accountability, perseverance, self-respect, and problem solving skills that have been logically linked to academic performance (these possible indirect effects are not depicted in figure 1.1, and are not tested in the current study).

Research domains and study questions

The study design includes assessments pertaining to each of the domains depicted in the logic model. Specifically, one set of domains, defined as primary confirmatory outcomes, pertains to the student-level social competence, behavior, and academic achievement outcomes depicted on the right side of figure 1.1. The other domain reflects intervention effects on the school environment (intermediate confirmatory outcomes).

Primary confirmatory research questions—student outcomes

Formally, the study is guided by three primary confirmatory research questions:

1. Do students in grades 4 and 5 who attend schools in the LIC intervention group exhibit higher scores on measures of academic achievement after two academic years of potential LIC exposure than their counterparts who attend schools in the control group?
2. Do students in grades 4 and 5 who attend schools in the LIC intervention group exhibit higher scores on measures of social competence after two academic years of potential LIC exposure than their counterparts who attend schools in the control group?
3. Do students in grades 4 and 5 who attend schools in the LIC program intervention group exhibit fewer problem behaviors after two academic years of potential LIC exposure than their counterparts who attend schools in the control group?

The analytic sample used to address research questions 1–3 comprises students in grade 4 and 5 intervention and control study classrooms in year 2 of the study who were also exposed to study teachers in year 1. Thus, the confirmatory research questions focus on the impacts of exposure to the LIC program over a two-year period. The LIC program is typically implemented schoolwide in all elementary grades. The design focuses on students with two years of exposure so that the impact estimates will more closely reflect the effectiveness of the program under conditions in which LIC would typically be implemented.

Although grade 2 and 3 students participated in the study, the primary analytic sample is restricted to grade 4 and 5 students so that outcomes based on student self-report data (which were not collected from grade 2 and 3 students) could be included (see chapter 2).

Two or three outcome measures are examined within each of the outcome domains. Two measures of academic achievement are examined—state standardized English language arts test scores and teacher reports of student academic competence. The social competence outcomes include teachers' appraisals of social skills and students' reports of altruism and empathy. Problem behavior measures include teacher reports of externalizing behavior and student reports of aggression and delinquent behavior. Multiple comparison procedures are used to adjust for the possibility of finding statistically significant results within each outcome domain due exclusively to chance factors.

Intermediate confirmatory research questions—school environment outcomes

As described above, the theory of action linking the LIC program to students' academic achievement, social competence, and problem behaviors holds that the program impacts students indirectly, via the school environment. The program is thought to positively affect students indirectly by promoting a caring environment and a common set of student behavioral expectations, and thereby by making the school environment more conducive to learning. To examine part of this logic model, the study estimates the impacts of the LIC program on two indicators of the school environment: student belongingness and school expectations.

Specifically, the following two intermediate confirmatory research questions are examined:

4. Do teachers and students in the LIC intervention group report greater levels of student belongingness after two years of program implementation than their counterparts in the control group?
5. Do teachers and students in the LIC intervention group report greater levels of school expectations consistent with character development after two years of program implementation than their counterparts in the control group?

Evidence supporting the logic model is provided if statistically significant positive impacts are found for research question 4 or 5. However, in no way is such evidence used to make inferences about the overall effectiveness of the intervention. Only research questions 1–3, involving student-level outcomes, are used to make inferences about the effectiveness of the program on measures within each outcome domain.

There are two measures for each of these outcome domains, one based on student survey data and the other based on teacher survey data. The analytic sample used to address research questions 4 and 5 comprises teachers of grades 2–5 and students in grade 4 and 5 study classrooms in year 2. Impacts are also examined on teacher reported outcomes for the grade 4–5 subsample of teachers—teachers in the same grades as students included in the analytic sample.

Exploratory research questions

To guard against false discoveries and to maintain adequate statistical power, the primary confirmatory research questions focused on a limited number of assessed outcome variables and subsamples. As noted above, inferences in this report about the effectiveness of the intervention

were made only with regards to the subsample and set of measures that were defined as confirmatory primary outcomes during the design stage of the study. Although the primary and intermediate questions focus on outcomes after two years of program implementation, we also gathered data on the key outcomes after the first year of program implementation. These data are used in exploratory analyses.

The purpose of these exploratory analyses was to obtain preliminary estimates of program impacts that could not be rigorously tested in the present study due to statistical power considerations.⁵ The results of the exploratory analyses were not used to make inferences about the overall effectiveness of the intervention but to provide descriptive information that can inform the confirmatory results and be rigorously tested in future studies. For example, depending on the pattern of results across years, the first year impact estimates can provide information about the possible role of cumulative exposure to LIC across years in affecting student outcomes, or the possibility of impacts diminishing after year 1. Because these analyses were exploratory, no corrections for multiple hypothesis testing were applied.

Specifically, the following exploratory research questions pertaining to year 1 program impacts were examined:

6. Do students in grades 4 and 5 who attend schools in the intervention group exhibit higher scores on measures of academic achievement and social competence, and fewer problem behaviors after one academic year of potential LIC exposure than their counterparts who attend schools in the control group?
7. Do teachers and students in the LIC intervention group report greater levels of student belongingness and school expectations after one year of program implementation than their counterparts in the control group?

The analytic sample used to address research question 6 comprised grade 4 and 5 students in study classrooms in year 1. Research question 7 was examined using data collected from teachers of grades 2–5, teachers in grades 4 and 5, and students in grade 4 and 5 study classrooms in year 1. The outcome measures for research questions 6 and 7 are identical to those used for research questions 1–5, except that the latter were assessed after one year of program implementation.

Organization of report

The remainder of the report is organized to provide a description of the methods used for the study, the data collected, and the findings from the implementation and impact analyses. Chapter 2 describes the study design, data collection, sample selection, and data analysis methods. It also describes the characteristics of study schools, participating teachers, and participating students as well as differences in baseline characteristics across intervention and control group schools. Chapter 3 describes the intervention, implementation fidelity, and differences in practices

⁵ If both year 1 and year 2 estimates were used to make inferences about program effectiveness, twice as many statistical tests would need to be controlled for to adjust for the chances of finding statistically significant results due exclusively to chance factors. Applying twice as many multiple hypothesis adjustments would reduce statistical power.

aligned with character education across intervention and control group schools. Chapter 4 presents the primary confirmatory and intermediate main impact findings. Chapter 5 describes the results of the exploratory analyses. And chapter 6 summarizes the key findings of the study.

2. Study design and methodology

The study uses a multisite, school-level randomized controlled trial to assess the impacts of the Lessons in Character (LIC) program on students' academic achievement, social competence, and problem behaviors. Fifty volunteer schools were randomly assigned to intervention and control conditions (25 schools per condition). All teachers of grades 2–5 in schools in the intervention condition (year 1 only) had the opportunity to attend a one-day training course. They also had the opportunity to receive approximately two hours of on-site coaching in year 1. Intervention group teachers were expected to incorporate 19–25 supplementary lessons into their classroom instruction for two academic years. No LIC professional development activities or coaching occurred during year 2.⁶ Teachers in control group schools, meanwhile, continued with their regular professional development activities and instructional practices.

Two cohorts of elementary schools participated in the intervention. Cohort 1 implementation took place in 2007/08 and 2008/09, with teacher professional development and coaching occurring in late summer and early fall of year 1 (2007). Cohort 2 schools were recruited in spring 2008, with implementation in 2008/09 and 2009/10. Both cohorts were pooled in the data analyses. The total study population consisted of 14,679 students in grades 2–5, served by 817 teachers, in 50 schools in California. As described in more detail later in this section, 4,683 students who were in grades 4 and 5 in year 2 were included in the primary impact analyses, and 5,674 students who were in grades 4 and 5 in year 1 were included in the exploratory analyses. Student outcomes were assessed using teacher reports on the Social Skills Rating System (SSRS) (Gresham and Elliott 1990), surveys of grade 4 and 5 students, and standardized English language arts achievement tests. Table 2.1 describes key design features of the study.

⁶ New teachers in intervention schools who joined the study in year 2 did not receive any training in the use of the curriculum, nor did they receive any on-site coaching support.

Table 2.1. Study characteristics and data collection schedule for Lessons in Character impact evaluation

| Characteristic | Description |
|--|--|
| Study design | Cluster-randomized trial |
| Unit of assignment | Schools |
| Sample characteristics | <p>14,679 students in 50 schools—schools in the intervention condition had the opportunity to implement Lessons in Character in grades 2–5 for two years.</p> <p>Primary analytic sample comprised of 4,683 4th and 5th grade students in the second implementation year.</p> <p>Exploratory analytic sample comprised of 5,674 4th and 5th grade students in the first implementation year.</p> |
| | 625 participating teachers in year 1. |
| | 550 participating teachers in year 2. |
| Statistical power estimates | For Type 1 error = .05, 80 percent or higher power to detect true impacts of 0.28 or larger for primary academic outcomes and 0.20 or larger for primary behavioral outcomes. |
| Implementation start date | Cohort 1: summer 2007; Cohort 2: summer 2008 |
| <i>Student measures</i> | |
| Standardized test scores | Cohort 1: spring 2007, 2008, 2009 |
| Primary: English language arts | Cohort 2: spring 2008, 2009, 2010 |
| Teacher-reported measures | Cohort 1: spring 2007, 2008, 2009 |
| Primary: social competence, externalizing behavior, and academic competence (SSRS) | Cohort 2: spring 2008, 2009, 2010 |
| Intermediate: student belongingness and school expectations | |
| Student-reported measures (grade 4/5) | Cohort 1: fall 2007; spring 2008 and 2009 |
| Primary: altruism, empathy, aggression, delinquency | Cohort 2: fall 2008; spring 2009 and 2010 |
| Intermediate: student belongingness and school expectations | |

Sampling and power estimates

To determine a priori the appropriate sample sizes required for the study, minimum detectable effect sizes (see Bloom 1995) were calculated based on the unit of randomization, the level of clustering, the availability of baseline explanatory variables, and other design characteristics, using the procedures described by Donner and Klar (2000), Murray (1998), Raudenbush (1997), and Schochet (2005).

A student attrition rate of about 25 percent was assumed for power estimation purposes, leaving 18 students per class at the end of the second implementation year available for analysis of the outcomes assessed. For the purposes of the power analyses, school-and teacher-level intraclass correlations of 0.15 were assumed for student academic performance and 0.07 for the student nonacademic primary outcomes (Hedges and Hedberg 2007; Schochet 2008; Murray and Hannan 1990; Murray and Blitstein 2003; Murray and Short 1996; Scheier et al. 2002). The intraclass correlations for classroom intermediate outcomes were expected to be 0.15. The statistical power analyses also assumed between- and within-school R^2 values of 0.50 (Schochet 2005). Finally, using a Bonferroni adjustment as a conservative approximation of the proposed resampling method (described below), the critical value of the statistical significance test (0.05) was divided by the number of primary measures analyzed within each outcome area (academic performance, social competence, and problem behaviors). For the intermediate outcome measures, the critical value was divided by the number of intermediate school environment outcomes (4).

With 25 schools per condition, 6 classrooms per school, and a minimum of 18 students in each classroom, the minimum detectable effect sizes were estimated at 0.27 standard deviations for student academic achievement and 0.20 for social competence and problem behaviors. The minimum detectable effect sizes for the intermediate outcomes were estimated at 0.42 for teacher-reported school environment measures and 0.21 for student-reported school environment measures. Further details related to estimation of minimum detectable effect sizes as well as statistical power estimation for the achieved sample are provided in appendix A.

Sample recruitment

Schools were recruited throughout California, using established WestEd marketing channels and mass mailings. Although the recruitment procedures differed for cohort 1 and cohort 2 schools, for both cohorts, recruitment efforts involved outreach to district superintendents and school principals and concluded with both schools and school districts signing a memorandum of understanding during the spring prior to the first program implementation year. Random assignment of schools was not conducted until after memoranda of understanding were received and baseline data were collected.

To be eligible to participate in the study, an elementary school had to serve students in grades 2–5, to have not implemented LIC in the past, and to have at least half of all teachers of grades 2–5 in each grade agree to participate. All traditional teachers of grades 2–5 and their students were eligible to participate. As noted above, only students in grades 4–5 were included in the primary and exploratory analytic samples. Special education teachers (that is, those who taught special

education classrooms exclusively) were excluded from the study sample, but special education students in traditional classrooms were eligible to participate.

Cohort 1 school recruitment

For cohort 1, messages describing the study and requesting participation were sent out to districts and schools using WestEd's Title IV and California Healthy Kids Survey electronic mailing lists. These mailing lists are maintained by the Health and Human Development Program housed at WestEd. A presentation was also made at the state Title IV county coordinator meeting in spring 2007. These activities resulted in approximately 102 expressions of interest from individual elementary schools.

Prescreening interviews with principals were conducted by the research team to screen out schools already implementing LIC, to ensure that school staff was fully aware of the requirements of participating in a randomized controlled trial, and to confirm that enough teachers were willing to participate in the study. During the prescreening interview, the principal was provided with an overview of the LIC program, the evaluation, and specific information about the process of obtaining parental consent. Also, attempts were made to schedule follow-up presentations at school faculty meetings.

Once oral confirmation of study participation was received, a memorandum of understanding was sent to each school and managing school district, outlining what support and possible compensation sites would receive for participating in the study, the roles and responsibilities of both research staff and site staff, and estimates of the time required to collect data.

The U.S. Office of Management and Budget approval for data collection was received on May 16, 2007—approximately four weeks prior to the end of the academic year for most of the schools in the recruitment pool. These four weeks were needed to complete recruiting and baseline data collection. The limited time available for recruitment, in combination with already full faculty meeting calendars, limited opportunities for the research team to present at faculty meetings and secure signed memorandums of understanding from school districts. Although 36 of the 102 schools that had expressed interest in participating in the study had informally agreed to participate in the study, formal approval was obtained from only 15 of them by the end of June 2007.

Cohort 2 school recruitment

A different process was used to recruit cohort 2 schools. Informational letters and statement-of-interest return facsimile forms were sent to principals in all elementary schools located in the five counties in the Los Angeles Metropolitan Statistical Area.⁷ Of the 2,226 principals to whom mailings were distributed, approximately 115 expressed an interest in participating in the study. Similar to the cohort 1 recruitment, prescreening interviews with principals were conducted and, if warranted, presentations at school faculty meetings were scheduled. Approximately 54 presentations at school faculty meetings were conducted to ensure that teachers understood the LIC program and the requirements for study participation. Of the 115 schools that expressed

⁷ The research team was based in Southern California. Recruitment targeted elementary schools in the Los Angeles area instead of the entire state of California to reduce costs associated with school recruitment site visits and implementation of the summer professional development trainings.

an interest in participating in the study, 29 had enough teachers in each grade who agreed to participate, provided formal approval, and participated in baseline data collection. In addition, a northern California district originally in the cohort 1 recruitment pool was re-contacted and recruited to participate in cohort 2, resulting in six additional cohort 2 schools. Overall, a total of 35 cohort 2 schools were recruited.

Incentives to participate

Schools and teachers were provided with several incentives to participate in the study. First, all participating schools received the LIC program materials and the opportunity to participate in the one-day program training at no cost. The market costs for the lesson materials provided ranged from \$16.33 (grade 2) to \$21.11 (grades 3–5) per student. The control group schools received the LIC materials and professional development services at the conclusion of the study. REL West also paid participating teachers \$100 a year (a maximum of \$300 over three years) as compensation for time spent completing surveys and other data collection activities. Within each school, a site coordinator was selected to assist in data collection activities (for example, distributing consent forms and packets to teachers, scheduling surveys, assisting with archival data collection). The site coordinator was paid \$500 over the three-year study period.

Participant consent and consent follow-up

The procedures used to secure parental permission for student study participation differed for cohort 1 and cohort 2 schools. For cohort 1, institutional review board requirements called for written parental permission (active parental consent) for students to participate in the study. Despite repeated efforts, however, significant difficulties were encountered in securing parental consent forms, with greater difficulty in low-performing schools with large percentages of racial/ethnic minorities. Midway through November 2007, parental consent forms had been returned by approximately 60 percent of eligible cohort 1 students, with 83 percent of returned consent forms indicating affirmative consent.

Because of these difficulties, a formal exemption from institutional review was requested for cohort 2 schools. The exemption was approved, recognizing that the study was investigating normal educational practices in a standard educational setting. As a result, passive consent procedures were used for cohort 2. Students in cohort 2 schools and their parents were notified of the study and given the chance to easily opt out if they preferred. Thus, active parental consent was required for cohort 1 students and passive consent was required for cohort 2 students. These differences in consent requirements resulted in substantial differences in student study participation rates across the two cohorts.

For cohort 1, materials requesting parental consent were distributed to site coordinators at each school in spring 2007 (prior to the first implementation year). Site coordinators, in turn, ensured that consent forms were distributed to all parents of students in grades 1–4. Each week, completed parental consent forms were express-mailed to REL West by the site coordinator. Two weeks after the consent materials were first distributed to parents, consent materials were redistributed to those parents who had not yet returned consent forms. Approximately one to two weeks after redistributing consent forms, baseline SSRS checklists for all students of consenting parents were distributed to teachers. In fall 2007, materials requesting parental consent were distributed to parents of newly enrolled students in grades 2–5 and to parents who had not returned consent forms during the prior spring. Baseline student surveys (grades 4 and 5 only)

were administered in schools approximately two to three weeks after consent forms were distributed in the fall.

To account for potential biases introduced by the nonreturn of consent forms in cohort 1, a nonresponse subsampling strategy was implemented in November 2007 to follow up with securing informed consent. Similar to the strategy used in the Moving to Opportunity (Orr et al. 2003) and DC Voucher (Wolf et al. 2007) evaluations, up to five students or parents among those who did not return informed consent forms were randomly sampled within each grade 2–5 classroom for more intensive follow-up. Teachers and administrators were provided with compensation in order to secure informed consent forms for these cases.⁸ No efforts were made to obtain parental consent for students who enrolled in cohort 1 schools after the fall semester of year 1. Approximately 590 students in grades 2–5 enrolled in cohort 1 schools after fall 2007. These students—14 percent of grade-eligible cohort 1 students enrolled—did not participate in data collection activities.

As described above, a passive consent procedure was used in most cohort 2 schools. A letter describing the study and procedures for opting out of the study was distributed to all parents of enrolled students in grades 1–4 in spring 2008 and to all parents of newly enrolled students in grades 2–5 in fall 2008. Because passive consent materials were not provided to parents of students who enrolled in cohort 2 schools after fall 2008, students who enrolled midway through the first implementation year or subsequently (2,167 students, or 14 percent of grade-eligible students) did not participate in data collection. Of the 11,639 students enrolled in cohort 2 schools in fall 2008, 66 had parents who refused to allow their participation in the study. Overall, 99 percent of cohort 2 students for whom consent letters were distributed were eligible to participate in data collection activities. More details are provided in the following discussion of sample selection and retention.

Random assignment

A key feature of the study design is the random assignment of schools to an intervention group that receives the LIC training and materials or to a control group that continues its regular professional development and instructional practices. As recruitment of schools moved into spring, schools were randomized in five waves, to allow sufficient time to schedule professional development training during the summer months. To improve precision of impact estimates and to guard against chance nonequivalence between randomly assigned conditions, schools were organized into blocks of four to six schools prior to randomization. Similar schools were stratified into groups based on their urban classification (nonrural or rural) and their level of academic performance (as indicated by California’s Academic Performance Index) in the year of random assignment. Table 2.2 describes the characteristics of the 12 randomization strata. Blocking prior to randomization was conducted only for the first two randomization waves. A series of dichotomous variables indicating block membership were included in the impact analysis models.

⁸ Further details about the subsampling strategy used and how recruited students in the subsample were handled in the analysis are provided later in this chapter (see Sensitivity analyses).

Table 2.2. Randomization strata characteristics

| Strata | Description | Schools | Teachers (year 1) | Students (year 1) |
|-----------------------------|------------------------------|---------|-------------------|-------------------|
| Cohort 1 | | | | |
| <i>Randomization wave 1</i> | | | | |
| 1 | Nonrural, low performing | 4 | 68 | 1,395 |
| 2 | Nonrural, high performing | 3 | 36 | 720 |
| 3 | Rural, low performing | 4 | 19 | 392 |
| 4 | Rural, high performing | 4 | 27 | 566 |
| Cohort 2 | | | | |
| <i>Randomization wave 2</i> | | | | |
| 5 | Nonrural, lowest performing | 4 | 76 | 1,786 |
| 6 | Nonrural, low performing | 6 | 94 | 2,264 |
| 7 | Nonrural, medium performing | 6 | 103 | 2,316 |
| 8 | Nonrural, high performing | 4 | 55 | 1,203 |
| 9 | Nonrural, highest performing | 5 | 73 | 1,658 |
| <i>Randomization wave 3</i> | | | | |
| 10 | Urban | 3 | 31 | 715 |
| <i>Randomization wave 4</i> | | | | |
| 11 | Mostly rural | 4 | 52 | 1,035 |
| <i>Randomization wave 5</i> | | | | |
| 12 | Rural/small town | 3 | 35 | 731 |
| Total | | 50 | 669 | 14,781 |

Source: California Department of Education 2009; U.S. Department of Education 2009.

Random assignment was conducted using the random number algorithm of the Stata 10.1 statistical package (StataCorp 2007). Half the schools were assigned to the intervention group, and half were assigned to the control group. Before schools were randomly assigned to intervention and control groups within blocks, blocks were randomly assigned to two different groups—one group in which the extra school in a block with an odd number of schools would be assigned to the intervention group (the “odd” group) and another group in which the extra schools would be assigned to the control group (the “even” group). Specifically, blocks were ranked by a randomly generated number, and every other block in the ranked sequence was allocated to the group in which the extra school would be assigned to the intervention group.

Schools were then randomly assigned to intervention and control groups within each block, based on the random number generation. Within each block, schools were ranked by a randomly generated number, and either every even-numbered school or every odd-numbered school in the sequence was assigned to the intervention group, depending on whether the block was randomly assigned to the odd or even group, as described above.

Sample selection and retention

School participation

As described above, 50 of the 217 schools that expressed interest in participating in the study provided formal approval, participated in baseline data collection activities, and subsequently were randomly assigned to intervention and control groups (see figure B1 in appendix B for a description of school and study participant recruitment and retention). All 50 schools continued

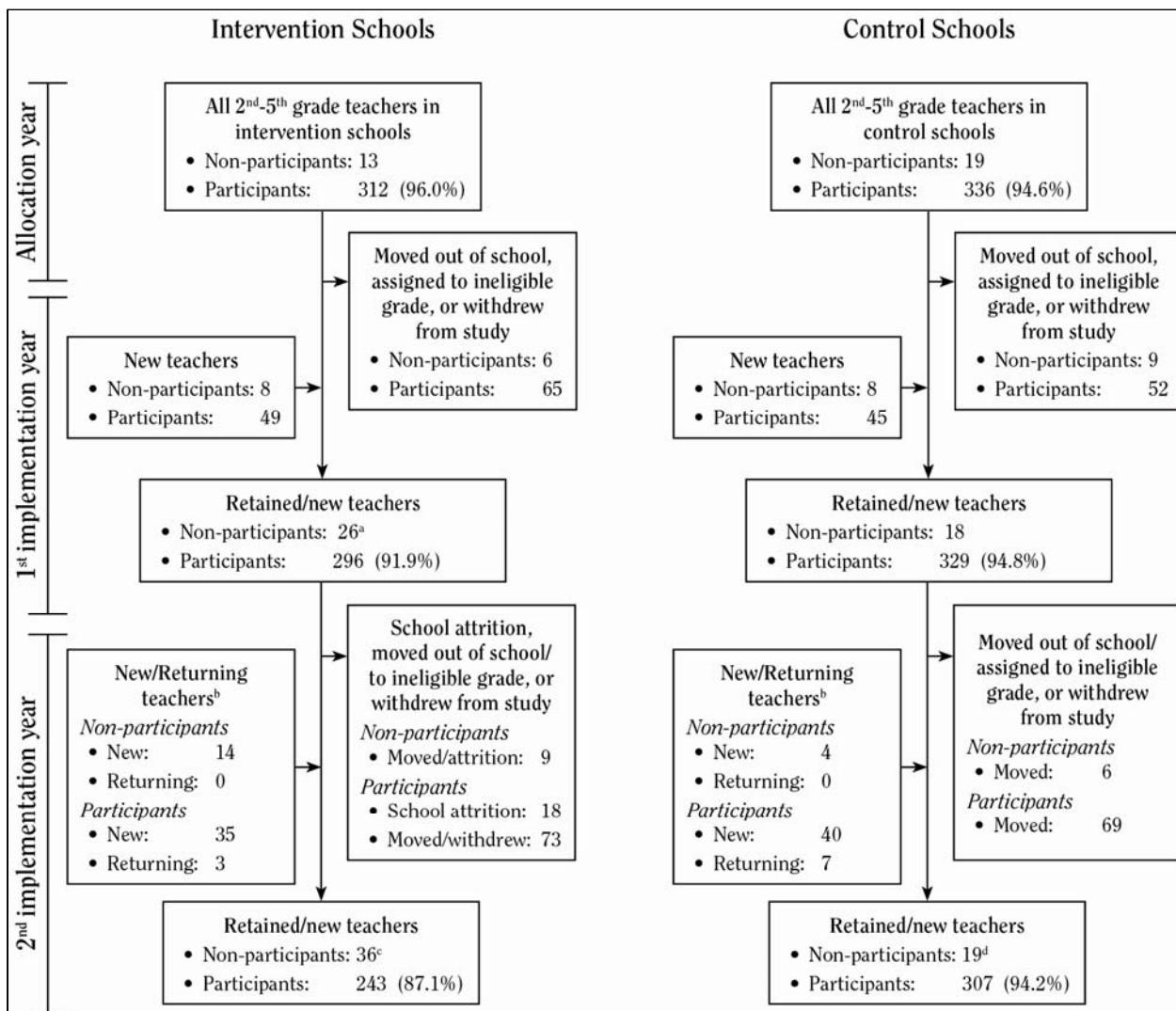
to participate during the first implementation year, although with varying degrees of involvement (see below). After the first implementation year, a small number of schools formally withdrew from the study and refused to participate in further data collection. Sensitivity analyses were conducted to ascertain how school attrition affected the impact estimates (see appendix G).

Teacher participation

The study targeted all general education teachers of grades 2–5 and their students in participating schools. As a condition for school participation, at least half of the teachers in each grade had to agree to participate in the study. Figure 2.1 and table 2.3 show details about the recruitment and mobility of teachers in participating schools prior to random assignment, during the first implementation year, and during the second implementation year (see figure B2 in appendix B for a grade-specific flow diagram of teacher recruitment and retention). Prior to random assignment, 95 percent of teachers (648 of 680) participated in the evaluation study, and 95 percent of teachers who agreed to participate provided baseline teacher-reported data. Pre-intervention participation rates were similar in the intervention and control groups.

In year 1, 117 teachers (18 percent) who participated in year 0 (pretest) were not retained. The majority of teachers not retained (101) were no longer employed at their schools during the first implementation year. Thirteen teachers who were still employed at their schools withdrew from the study after random assignment, with the majority of withdrawals taking place in intervention schools. Intervention and control group school differences in teacher retention and withdrawal rates were not statistically significant at the 0.05 level.

Figure 2.1. Flow diagram of nonparticipating and participating teachers into and out of study schools



Note: Numbers in parentheses are the percentage of eligible teachers participating in the study in the spring of the academic year.

a. Includes 11 teachers who withdrew consent to participate in the study.

b. Returning teachers were those who moved out of school after the allocation year or the fall of the first implementation year but returned by the spring of the second implementation year.

c. Includes six teachers who withdrew consent to participate in the study.

d. Includes three teachers who withdrew consent to participate in the study.

Newly employed teachers were recruited to participate in the study during both implementation years. In year 1, 110 new grade-eligible teachers were hired, of whom 94 (85 percent) provided affirmative consent to participate in the study. By the spring of year 1, 92 percent of grade-eligible teachers in intervention group schools and 95 percent of grade-eligible teachers in control group schools actively participated in the study.

As shown in table 2.3, several factors combined to reduce teacher retention more significantly in intervention group schools (69 percent) than in control group schools (79 percent) in year 2, although differences between intervention and control group schools in the retention of year 1 teachers and teacher mobility are not statistically significant. Also in year 2, 85 new teachers were hired and 75 agreed to participate in the study (88%). The 35 newly hired teachers in intervention schools did not receive training before having the opportunity to implement the LIC curriculum.

In the spring of year 2, 87 percent of grade-eligible teachers in intervention group schools and 94 percent of grade-eligible teachers in control group schools participated in the study. Although year 2 differences in study participation across intervention and control group schools were not statistically significant, these differences in teacher participation rates had implications for group differences in student study participation, as described below.

Table 2.3. Teacher participation and mobility rates of participating teachers by experimental group

| Participation or mobility measure | Intervention | Control | Difference | z-test ^a | p-value |
|--|--------------|---------|------------|---------------------|---------|
| <i>Year 0 – Spring</i> | | | | | |
| Participation rate of year 0 teachers ^b | 96.00 | 95.18 | 0.82 | 0.29 | 0.77 |
| <i>Year 1 – Spring</i> | | | | | |
| Retention of year 0 teachers ^c | 79.17 | 84.52 | -5.35 | -1.77 | 0.08 |
| Exited school/grade (year 0 teachers) ^d | 17.31 | 14.88 | 2.43 | 0.84 | 0.40 |
| Withdrew from study (year 0 teachers) ^e | 3.53 | 0.60 | 2.93 | 1.40 | 0.16 |
| Rate of entering teachers ^f | 16.55 | 13.68 | 2.87 | 0.87 | 0.38 |
| Participation rate of year 1 teachers ^b | 91.93 | 94.81 | -2.88 | -0.43 | 0.67 |
| <i>Year 2 – Spring</i> | | | | | |
| Retention of year 1 teachers ^c | 69.26 | 79.03 | -9.77 | -1.63 | 0.10 |
| School drop out | 5.47 | 0.00 | 5.47 | — | — |
| Exited school/grade (year 1 teachers) ^d | 24.10 | 20.06 | 4.04 | 1.18 | 0.24 |
| Withdrew from study (year 1 teachers) ^e | 2.03 | 0.91 | 1.12 | 0.52 | 0.60 |
| Rate of entering teachers ^f | 15.64 | 15.31 | 0.33 | 0.16 | 0.87 |
| Participation rate of year 2 teachers ^b | 87.10 | 94.17 | -7.07 | -1.17 | 0.24 |

— is not applicable.

a. Test statistic is based on a multilevel regression model that accounted for nesting of teachers within schools. No multiple comparison adjustments were applied.

b. Number of participating teachers divided by the number of eligible teachers (multiplied by 100).

c. Number of participating teachers retained in the study from the previous academic year divided by the number of participating teachers in the previous academic year (multiplied by 100).

d. Number of participating teachers who exited the school or eligible grade divided by the number of participating teachers in the previous academic year (multiplied by 100).

e. Number of teachers who withdrew from the study divided by the number of participating teachers in the previous academic year (multiplied by 100).

f. Number of new participating teachers divided by the number of eligible teachers (multiplied by 100).

Source: Teacher roster data files.

Student participation

Student study participation was largely determined by three factors: the year and semester that the student was enrolled in the school, parental consent, and enrollment in participating teachers' classrooms. As discussed above, parental consent forms were distributed to parents of enrolled students in grades 1–4 in the spring of year 0 and to parents of newly enrolled students in grades 2–5 in the fall of year 1.⁹ Students who transferred into study schools after the fall of year 1 were not eligible to participate because consent forms were not distributed to parents of students who appeared on school rosters for the first time in the spring of year 1 or subsequently. Additionally, students who transferred out of study schools after the fall of year 1 were not retained in the analysis.

The requirement of active parental consent for cohort 1 students limited participation to approximately 61 percent of grade-eligible students in year 1. However, because the majority of eligible students attended cohort 2 schools, which used a passive parental consent requirement, only 8 percent of the total potential year 1 student sample was barred from participation due to parental consent requirements. Finally, only students with parental consent who were enrolled in participating teachers' classrooms in year 1 or year 2 were eligible to participate in data collection. This is because teacher-reported student outcome data were collected only in participating teachers' classrooms. Teachers who did not agree to participate in the study were mostly not amenable to completing SSRS forms in their classrooms. Thus, the level of student study participation is directly linked to the level of teacher participation.

Figure 2.2 and table 2.4 depict how the analytic samples were determined. The primary analytic sample consisted of students with parent consent who remained at the study school and were enrolled in a participating teacher's classroom at the time of the spring year 2 data collection. Because the primary confirmatory research questions focus on grade 4 and 5 students exposed to participating teachers during both implementation years, study participation patterns are shown for students in grades 3 and 4 in year 1 because these students were grade-eligible for two years of LIC exposure. Overall, 37 percent of students enrolled in grades 3 and 4 in study schools in year 1 (2,725) were excluded from the primary analytic sample in the spring of year 2, primarily due to student mobility out of study schools (17 percent) or enrollment in nonparticipating classrooms (12 percent). Approximately 8 percent (584) were excluded because they did not have parental permission to participate. Although overall primary analytic sample participation rates were lower in intervention group schools (58 percent) than in control group schools (68 percent), the difference was not statistically significant. Rates of parental consent, enrollment in participating classrooms, and mobility out of study schools were also not statistically different among students eligible for the primary sample in intervention and control group schools.

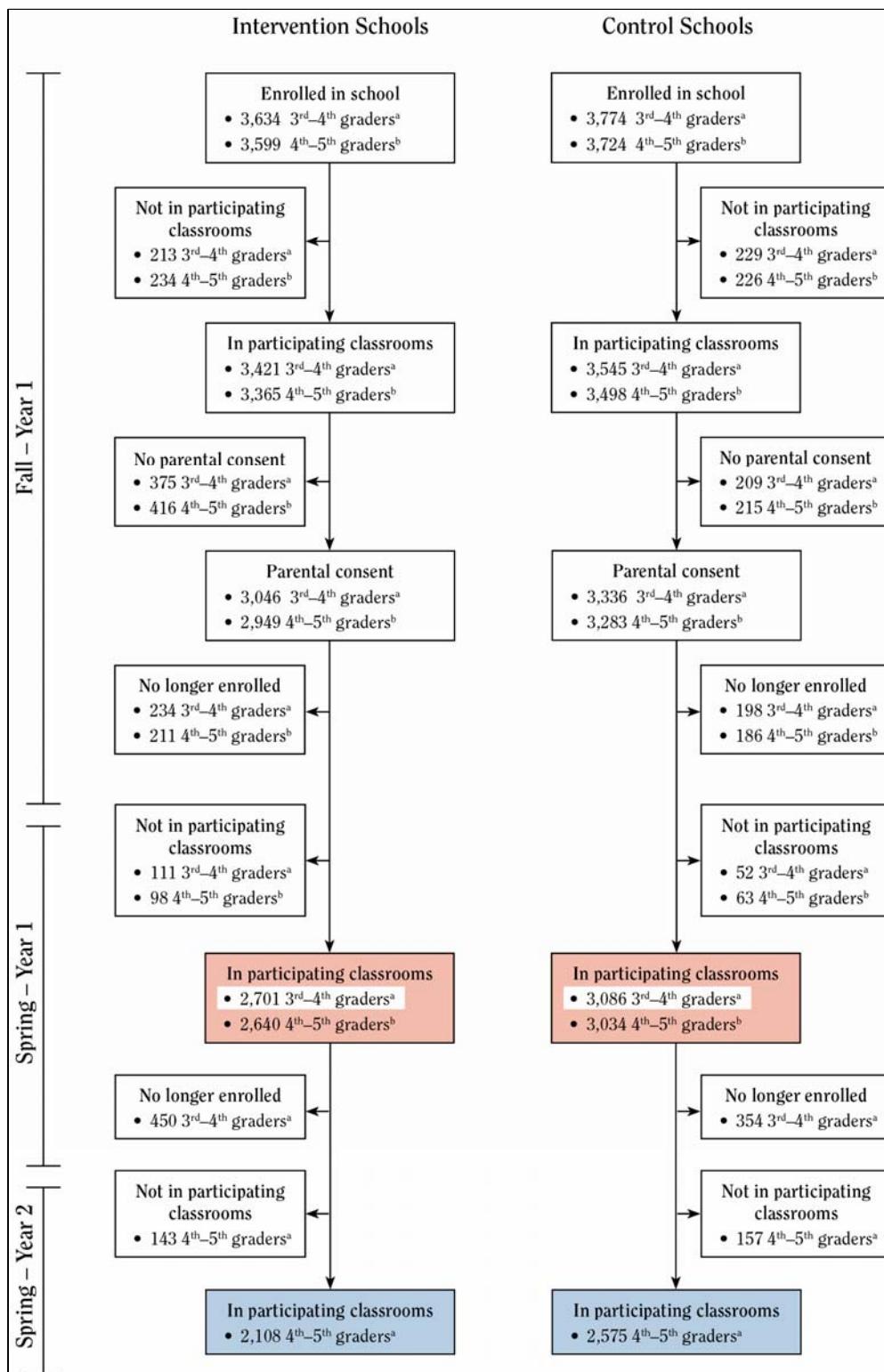
Fewer grade-eligible students were excluded from the exploratory analytic sample than from the primary sample because students eligible to be in the exploratory analytic sample were required to be enrolled for only one academic year. Approximately 23 percent of grade 4 and 5 students enrolled in study schools in the fall of year 1 (1,649) were excluded from the exploratory analytic sample in spring of year 1. Approximately 9 percent (631) were excluded from the

⁹ Consent forms were redistributed in the fall of year 1 to cohort 1 parents who did not return consent forms the previous spring. Additionally, extensive follow-up efforts were conducted to obtain consent forms from a subsample of cohort 1 parents who did not return consent forms by November 2008

sample because they did not have positive parental consent, 8 percent (621) were excluded because they were enrolled in a nonparticipating classroom, and 5 percent were excluded because they were no longer enrolled in the school in the spring semester. Overall, approximately 73 percent of grade 4 and 5 students enrolled in intervention group schools in the fall of year 1 were included in the exploratory analysis sample, compared with 81 percent of students in control group schools. Differences between intervention and control group schools in participation rates, parental consent, enrollment in participating classrooms, and mobility out of study schools were not statistically significant among students eligible for the exploratory analytic sample.

Figure B3 in appendix B provides a more detailed description of participation by student grade prior to random assignment, during the first implementation year, and during the second implementation year among all students in all grades targeted for data collection activities.

Figure 2.2. Flow diagram detailing student analytic samples



Note: The **blue box** represents the primary analytic sample, and the **red box** represents the exploratory analytic samples.

a. Students eligible to be in the primary analytic sample.

b. Students eligible to be in the exploratory analytic sample.

Table 2.4. Student participation rates by sample inclusion criterion and experimental group

| Criterion | Intervention group | Control group | Difference | z-test ^a | p-value |
|---|--------------------|---------------|------------|---------------------|---------|
| Primary analytic sample (grades 3 and 4 in year 1; grades 4 and 5 in year 2) | | | | | |
| <i>Year 1—fall</i> | | | | | |
| Enrolled in school | 100.00 | 100.00 | | | |
| —in participating classroom ^b | 94.14 | 93.13 | 1.01 | 0.46 | 0.64 |
| —with parental consent ^c | 83.82 | 88.39 | -4.57 | -0.23 | 0.82 |
| <i>Year 1—spring</i> | | | | | |
| —continued enrollment ^d | 77.38 | 83.15 | -5.77 | -0.48 | 0.63 |
| —in participating classroom ^e | 74.33 | 81.77 | -7.44 | -0.77 | 0.44 |
| <i>Year 2—spring</i> | | | | | |
| —continued enrollment ^f | 61.94 | 72.39 | -10.45 | -1.12 | 0.26 |
| —in participating classroom ^g | 58.01 | 68.23 | -10.22 | -0.86 | 0.38 |
| Exploratory analytic sample(grades 4 and 5 in year 1) | | | | | |
| <i>Year 1—fall</i> | | | | | |
| Enrolled in school | 100.00 | 100.00 | | | |
| —in participating classroom ^b | 93.50 | 93.93 | -0.43 | -0.39 | 0.69 |
| —parental consent ^c | 81.94 | 88.16 | -6.22 | -0.47 | 0.63 |
| <i>Year 1—spring</i> | | | | | |
| —continued enrollment ^d | 76.08 | 83.16 | -7.08 | -0.51 | 0.61 |
| —in participating classroom ^e | 73.35 | 81.47 | -8.12 | -0.45 | 0.65 |

a. Test statistic is based on multilevel regression model that accounted for nesting of students within teachers and teachers within schools. No multiple comparison adjustments were applied.

b. Number of students in participating classrooms in fall year 1 divided by the number of students enrolled in fall of year 1 (multiplied by 100).

c. Number of students in participating classrooms in fall year 1 with parental consent divided by the number of students enrolled in fall of year 1 (multiplied by 100).

d. Number of students in participating classrooms in fall year 1 with parental consent who remained enrolled in spring year 1 divided by the number of students enrolled in fall of year 1 (multiplied by 100).

e. Number of students in participating classrooms in fall year 1 with parental consent who remained enrolled and in participating classrooms in spring year 1 divided by the number of students enrolled in fall of year 1 (multiplied by 100).

f. Number of students in participating classrooms in fall year 1 with parental consent who remained enrolled through spring of year 2 and in participating classrooms in spring year 1 divided by the number of students enrolled in fall of year 1 (multiplied by 100).

g. Number of students in participating classrooms in fall year 1 with parental consent who remained enrolled and in participating classrooms through spring of year 2 divided by the number of students enrolled in fall of year 1 (multiplied by 100).

Source: Student roster data.

Teacher and student migration and crossovers

A very small number of teachers transferred to different schools between the spring of year 0 and the fall of year 1, all moving from intervention schools to control schools. To maintain statistical equivalence between intervention and control schools, these teachers were included in their originally assigned condition in the analysis. No other teachers migrated across participating schools during the study period.

During study implementation, 88 students in the primary analytic sample and 78 in the exploratory analytic sample migrated across study schools (table 2.5). Of those who migrated across study schools, approximately two-thirds moved from a control school to an intervention school or vice versa. In the primary sample, 42 students moved from a control school to an intervention school, and 18 moved from an intervention school to a control school. Similar

differences were apparent for the exploratory sample—46 students moved to an intervention school, and 13 moved to a control school. Despite their apparent magnitude, differences across intervention and control schools in the percentages of migrant students who were initially enrolled in an intervention school but subsequently moved to a control school (and vice versa) were not statistically significant ($p=0.25-0.28$). To maintain the integrity of random assignment, crossover students were analyzed in their original assigned condition. Sensitivity analyses showing the consequences of including migrant and crossover students for the impact estimates are reported in appendixes G and H.

Table 2.5. Students who migrated across schools by crossover status

| Condition | Primary sample | | Exploratory sample | |
|---|-----------------------|-------------------|---------------------------|-------------------|
| | Number | Percent | Number | Percent |
| <i>Noncrossovers</i> | | | | |
| Transferred across intervention schools | 13 | 14.6 | 10 | 11.4 |
| Transferred across control schools | 16 | 18.0 | 9 | 10.2 |
| <i>Crossovers</i> | | | | |
| Transferred from intervention to control school | 18 | 20.2 ^a | 13 | 14.8 ^a |
| Transferred from control to intervention school | 42 | 47.2 ^a | 46 | 52.3 ^a |
| <i>Total</i> | 89 | 100.0 | 78 | 100.0 |

Note: Student migrants refer to students who enrolled in two or more study schools during the study implementation period. Crossovers refer to students who were initially enrolled in an intervention school (control school) but subsequently moved to a control school (intervention school).

a. A two-tailed test that adjusted for intraclass correlations was used to test intervention/control school differences in the proportion of migrants who were classified as crossovers. No multiple comparison adjustments were applied. No differences were statistically significant at the 0.05 level.

Source: Student roster data.

Data collection

The evaluation relies on school archival data, grade 4 and grade 5 student surveys, teacher surveys, principal interviews, and site observations to measure student outcomes, the school environment, and implementation fidelity. Table 2.6 describes the outcome measures collected. Except for survey items asking specific questions about LIC implementation, all measures were collected in both intervention and control group schools. Tables 2.7 and 2.8 show the data collection schedule. Appendix C provides a detailed description of the items used to assess each outcome measure used in the primary and exploratory analyses. Appendix D presents internal consistency reliabilities (Cronbach's alpha) and intraclass correlations for each outcome measure, by analytic sample.

Table 2.6. Measurement matrix of primary and intermediate outcome variables

| Construct | Items | Source | Alpha reliability | Reference |
|--|-------|----------------|-------------------|--|
| Student outcomes | | | | |
| <i>Academic achievement</i> | | | | |
| English language arts tests ^a | 65–75 | School records | 0.93–0.94 | California Department of Education (2010a) |
| Academic competence ^a | 4 | Teacher | 0.90 | Gresham and Elliott (1990) |
| <i>Social competence</i> | | | | |
| Social Skills Rating System ^a | 30 | Teacher | 0.94 | Gresham and Elliott (1990) |
| Altruism ^a | 5 | Student | 0.86 | Characterplus (2002) |
| Empathy ^a | 11 | Student | 0.72 | Funk et al. (2003) |
| <i>Problem behaviors</i> | | | | |
| Externalizing ^a | 6 | Teacher | 0.88 | Gresham and Elliott (1990) |
| Aggression ^a | 6 | Student | 0.83 | Orpinas and Frankowski (2001) |
| Delinquent behavior ^a | 7 | Student | 0.71 | Kisker et al. (2004) |
| School environment/teacher outcomes | | | | |
| <i>School environment</i> | | | | |
| School expectations ^b | 5 | Student | 0.87 | Characterplus (2002) |
| School expectations ^b | 5 | Teacher | 0.94 | Characterplus (2002) |
| Students' feelings of belonging ^b | 12 | Student | 0.87 | Characterplus (2002) |
| Students' feelings of belonging ^b | 12 | Teacher | 0.89 | Characterplus (2002) |

a. Primary and exploratory outcomes.

b. Intermediate confirmatory outcomes.

Table 2.7. Data collection schedule

| Measure | Year 0 | Year 1 | Year 2 |
|---|--------|----------------------|-----------------------------------|
| <i>Student outcome measures</i> | | | |
| English language arts standardized tests | | | |
| Student surveys | Spring | Spring | Spring |
| Teacher SSRS | Spring | Fall/spring | Fall ^a /spring |
| <i>Teacher practice/fidelity measures</i> | | | |
| Teacher surveys | Spring | Spring | Spring |
| Teacher implementation logs (intervention only) | | Winter/spring/summer | Winter/spring/summer ^b |
| Principal interviews | | Spring | Spring |
| School site observation checklist | | Spring | Spring |

a. Fall student survey was collected from cohort 1 students only.

b. Implementation logs were collected from cohort 2 teachers in only the winter and the summer of year 2.

Table 2.8. Data collected by source, school grade, and year

| Year 0 (2006/07 for cohort 1, 2007/08 for cohort 2) | | | Year 1 (2007/08 for cohort 1, 2008/09 for cohort 2) | | | Year 2 (2008/09 for cohort 1, 2009/10 for cohort 2) | | |
|---|------|--------|---|--------|------|---|--|--|
| | Fall | Spring | Fall | Spring | Fall | Spring | | |
| <i>Grade 1</i> | | | | | | | | |
| SSRS | | + | | | | | | |
| Survey | | | | | | | | |
| Test scores | | | | | | | | |
| <i>Grade 2</i> | | | | | | | | |
| SSRS | | + | | + | | | | |
| Survey | | | | | | | | |
| Test scores | | + | | + | | | | |
| <i>Grade 3</i> | | | | | | | | |
| SSRS | | + | | + | | | | |
| Survey | | | | | | | | |
| Test scores | | + | | + | | | | |
| <i>Grade 4</i> | | | | | | | | |
| SSRS | | + | | | | | | |
| Survey | | | | | | | | |
| Test scores | | + | | + | | | | |
| <i>Grade 5</i> | | | | | | | | |
| SSRS | | | | | | | | |
| Survey | | | | | | | | |
| Test scores | | | | | | | | |

+ Measures collected.

Note: The represents the primary analytic sample, and the represents the exploratory analytic samples.

a. Fall student survey was collected from cohort 1 students only in year 2.

Primary and intermediate outcome data collection and measures

State English language arts assessments. Students' English language arts achievement data from district-administered, state-mandated standardized tests for all students were collected for the years before and during exposure to the study teachers. California administers the California Standards Tests (CST)—an assessment that is criterion-referenced to state standards—to students in grades 2–5. As with other state tests, all questions on the CST are evaluated by committees of content experts, including teachers and administrators, to ensure their appropriateness for measuring the California academic content standards. In addition to content review, all items are reviewed and approved to ensure their adherence to the principles of fairness and to ensure that no bias exists with regard to characteristics such as gender, ethnicity, and language. The English language arts CST includes 65–75 multiple-choice items, depending on the grade-level tests, that assess five performance clusters: word analysis and vocabulary, reading comprehension, literary response and analysis, written conventions, and writing strategies. The total scale score across these five domains is used as an outcome in the analysis. Reported reliability figures for the English language arts CST range from 0.93 to 0.94 (California Department of Education 2010a).

The CST is not vertically scaled and thus does not have the same meaning across different grade levels. So that test score data from all of the grades can be analyzed together, test score data were normalized within each grade by subtracting the state mean from each student's score and dividing by the state standard deviation. Normalized in this way, the test score data represent the relative ranking of students in the analytic sample rather than the absolute level of performance,

and the impact estimates (see below) reflect the impacts relative to the distribution of scores in the state. Because the CST does not measure the same academic skills across different grade levels, scores cannot be used to precisely assess growth in performance across time. As shown in table 2.8, student test score data were collected for grades 2–4 in year 0 (baseline), grades 2–5 in year 1, and grades 3–5 in year 2. The English language arts CST scale score was used as a primary confirmatory outcome to address research question 1 and as an exploratory outcome to address research question 6.

Social Skills Rating System teacher reports. Teacher reports on the SSRS (Gresham and Elliott 1990) were used to assess student social skills, problem behaviors, and academic competence. The SSRS forms were distributed to onsite coordinators, who in turn distributed the forms to participating teachers. Teachers were instructed to place completed forms in a provided sealed envelope and to give the sealed envelope to the site coordinator. The site coordinator sent teachers' completed packets to WestEd in postage-paid express mail envelopes. Teacher follow-up occurred at biweekly intervals.

The SSRS-Teacher is a 57-item multidimensional instrument assessing student social and academic functioning. Across constructs assessed by the instrument, internal consistency reliabilities (alphas) range from 0.78 to 0.94; test/retest reliabilities range from 0.75 to 0.93 (Gresham and Elliott 1990). The elementary SSRS-Teacher assesses the subdomains of cooperation, assertion, self-control, externalizing problems, internalizing problems, hyperactivity, and academic competence. The total scores for academic competence, social competence, and externalizing behavior were used to evaluate research questions 1, 2, and 3, respectively, as well as research question 6. Each of these scores was coded so that higher values reflect higher levels on the named measure. For example, higher scores on the social skills scale score correspond to higher levels of social skills. Higher scores on the externalizing scale reflect higher levels of problem behavior. Sample-specific alphas ranged from 0.76 (academic competence) to 0.97 (social skills). Appendix D provides more detailed information about study-specific internal consistency reliabilities. SSRS reports were collected for students in grades 1–4 in the spring of year 0 (baseline), for students in grades 2–5 in the spring of year 1, and students in grades 3–5 in the spring of year 2.

Student survey. A 35-minute survey assessing behaviors, attitudes, and values consistent with the goals of character education was administered to all grade 4 and grade 5 students during the fall and spring semesters of the first implementation year and the spring semester of the second implementation year. Surveys were administered by experienced proctors to students with parental consent during a single class period.

Using items and subscales from existing instruments, the survey assessed student empathy (Funk et al. 2003), altruism (Characterplus 2002), school engagement and disengagement (Furrer and Skinner 2003), aggression (Orpinas and Frankowski 2001), delinquent behavior (Kisker et al. 2004), autonomy/influence, school expectations, and students' feelings of belonging (Characterplus 2002). The student empathy ($\alpha = 0.72$ – 0.77), altruism ($\alpha = 0.79$ – 0.93), aggression ($\alpha = 0.82$ – 0.84), and delinquent behavior ($\alpha = 0.74$ – 0.75) subscales were used in the confirmatory primary impact analyses (research questions 2 and 3) and the

exploratory analyses (research question 6).¹⁰ Student reports of school expectations ($\alpha = 0.64$ – 0.87) and feelings of belonging ($\alpha = 0.85$ – 0.88) were used in the confirmatory analyses of intermediate classroom outcomes (research questions 4 and 5) and in the exploratory analyses (research question 7). Appendix D provides detailed information about sample-specific alpha estimates. Similar to the SSRS-based measures, scores on each of the scales were constructed so that higher values reflect higher levels on the named measure (that is, students with high scores on the delinquent behavior scale reported that they engaged in delinquent behavior more frequently than students with low scores).

Teacher surveys. Using the same procedure as described for the SSRS, all intervention and control group teachers were surveyed prior to random assignment during the spring of year 0 and in the spring of year 1 and year 2. Surveys were also administered to new teachers who entered participating schools and agreed to participate in the study after year 0. These surveys were administered as soon as the research team became aware of the new teacher, usually when rosters were collected in the fall and spring.

The surveys assessed activities related to social and character development that were implemented in the classroom and school and classroom climate (students' feelings of belonging, school expectations, parent and staff relations, staff culture of belonging, school leadership, and character development modeling). The teacher survey also contained questions about additional professional development that the teachers participated in during the implementation year as well as questions assessing teachers' communication and collaboration. The items on the survey assessing students' feelings of belonging ($\alpha = 0.95$ – 0.97) and school expectations ($\alpha = 0.90$ – 0.91) were used in the confirmatory analysis of intermediate classroom outcomes (research questions 4 and 5) and in the exploratory analysis (research question 7).

Implementation data collection and measures

Attendance logs. Teacher attendance data collected at the trainings were used to describe the percentage of teachers in intervention group schools who participated in the professional development. These data were also used to document attendance of school administrators at the trainings.

Teacher implementation logs. Teachers in schools in the intervention group filled out logs that assessed the frequency with which LIC, Daily Oral Language with Character, and Writing with Character lessons were delivered as well as teachers' perceptions of the extent that the lessons were delivered as described in the LIC implementation guides. These logs were used to assess the amount of the program delivered to students in intervention group schools. One LIC core lesson implementation log and two to three Daily Oral Language with Character and Writing with Character lesson logs (covering different calendar periods) were distributed to each applicable intervention group teacher at the beginning of the academic year. Implementation logs were collected (via facsimile) from cohort 1 intervention group teachers three times over the academic year (December, March, and May or June) in year 1 and year 2. The same data collection schedule was used in year 1 for cohort 2. To ease teacher burden, implementation data

¹⁰ Cronbach alphas are presented in the form of ranges across samples (confirmatory, exploratory) and measurement points (baseline, post-test). See appendix D for details.

were collected from cohort 2 teachers in December and May or June exclusively in year 2.¹¹ The LIC lesson logs ask teachers to write the date they implemented specific lessons. At each data collection point, teachers were instructed to fax the entire completed and non-completed LIC lesson log to the research team.¹² If dates were not listed next to applicable lessons or logs were not received, teachers were assumed to have not implemented the lessons. The Daily Oral Language with Character and Writing with Character logs ask teachers to indicate the approximate number of lessons (0, 1-5, 6-10, etc.) implemented during specific time periods (i.e., September to December, January to March, April to May, or January to May), depending on the reference period for the implementation log. Responses to these questions were summed across each data collection point to estimate the number of Daily Oral Language with Character and Writing with Character lessons implemented during the academic year. Responses to the lesson logs, particularly for the Daily Oral Language with Character and Writing with Character logs, could have been adversely affected by recall errors.

Teacher surveys. Teacher survey questions about professional development and activities aligned with character education were used to describe program implementation and differences in activities between intervention and control group classrooms. The professional development questions ask about participation in professional development activities during the 12-month period prior to survey completion. The questions about activities aligned with character education ask about implementation in the 6 months prior to survey completion. These questions were included on the baseline, year 1, and year 2 surveys.

Site observations. A site observation checklist was developed to assess evidence of social and character development activities in each school. This instrument was based on a similar instrument used in the Social and Character Development Impact Evaluation (Kisker et al. 2004). Evidence of posters and other artifacts indicative of social and character development activities was documented in each grade 4 and 5 classroom as well as in the public areas of the school. This observation checklist was used to document character education-like activities in both intervention and control group schools. Site observations were conducted by trained proctors during the student survey administration site visit.

Baseline data collection and measures

Specific pre-intervention covariates were selected based on study design considerations and judgments about the extent to which they were expected to explain variance in outcome variables. The following pre-intervention covariates were included in the models examining outcomes. Details regarding which covariates were included in specific models are provided in appendix F.

Pretest measures of outcomes. Pretest measures of each outcome variable were used to assess baseline equivalence and as covariates in the relevant impact analyses. School-level pretest measures (that is, the mean of pretest variables in the school) were also included as covariates in

¹¹ In year 2, the implementation log for cohort 2 was modified so that teachers reported on Daily Oral Language with Character and Writing with Character implementation during the January through May/June timeframe.

¹² The LIC lesson log thus provides a cumulative record of lesson implementation in the sense that if respondents only returned their log in June – complete data on LIC lesson implementation during the entire academic year would be available if respondents filled in the dates for all the LIC lessons delivered.

the impact analysis models. Moreover, because surveys were not administered to grade 3 students, student-level pre-intervention survey measures were not available for students who were in grade 4 in year 2. Therefore, only school means of pretest grade 4 and 5 student survey measures—not student-level pretest measures—were used as covariates in impact analysis models when outcomes based on student surveys were examined.

Demographic variables. Student and demographic information was obtained from district archival records. Variables obtained include student grade in school, gender, race/ethnicity (non-Hispanic White, Asian, Black, Hispanic, other), English language learner status, special education status, student birth date, and parent education level (no high school degree, high school degree, some college, college degree or more). The following demographic variables were used as covariates in the student impact analysis models and to assess baseline equivalence of intervention and control group schools: grade in school, gender, race/ethnicity, English language learner status, and parent education level. Teacher demographic information was obtained from the teacher surveys administered to teachers each spring.

Teacher surveys. As described above, intervention and control group teachers were surveyed prior to random assignment during the spring of year 0. In addition to using pretest measures of each teacher outcome measure to assess baseline equivalence and as a covariate in the applicable impact analysis model, pretest measures of teachers' years employed at the school, grade taught, gender, and race/ethnicity were also used for these purposes. Teacher pre-intervention reports of activities related to social and character development that were implemented in the classroom were also used to assess baseline equivalence.

Design variables. Because schools were randomized to conditions within strata, dichotomous variables for all but one stratum were included as covariates in the impact analysis models. The inclusion of randomization strata ensures that the impact analysis model conforms to the strategy used in random assignment by adjusting the degrees of freedom to reflect the number of randomizations performed. Its inclusion also potentially explains variance in the outcome variables, thereby increasing precision of estimated impacts. Because randomization strata are nested within each cohort, a dichotomous indicator for cohort membership is not included in the impact analysis models.

Other school-level variables. School-level academic performance and an index reflecting school demographic characteristics were also used to assess group equivalence and included as covariates in the impact analyses. School-level academic performance was assessed using the Academic Performance Index, a summary measure of school academic performance developed and used by the California Department of Education. The index is based on standardized achievement tests. The 2007 base score was used for cohort 1 schools, and the comparable 2008 growth index was used for cohort 2 schools.¹³ The index is publicly available for most schools in the state via the California Department of Education's website (California Department of Education, 2011).

¹³ The California Department of Education designed the Academic Performance Index reporting system so that the “base score” for a particular year is based on the same tests as the “growth score” for the subsequent year (California Department of Education 2010c).

Response rates

Teacher-level data—all eligible teachers

Table 2.9 shows teacher and student response rates by data source for the overall sample and for intervention and control group schools. Among all eligible teachers, overall teacher survey response rates ranged from 82 percent to 90 percent. Differences between intervention and control group schools in baseline, year 1, and year 2 survey response rates are not statistically significant at the 0.05 level. Although not statistically significant at conventional levels, year 2 teachers in intervention group schools were less likely to return surveys than their counterparts in control group schools, with a difference of 20 percentage points (71 percent compared with 91 percent). A difference of this magnitude in response rates could bias estimates of program impacts. To assess the potential for attrition-related bias, analyses of intervention and control group school differences in the retained sample were conducted and are described in the following section.

Teacher-level data—eligible teachers in grades 4-5

Because the primary and exploratory student analytical samples are restricted to students in grades 4-5, response rates are also shown for teachers serving students in the analytic sample – those serving grades 4-5. Response rates ranged from 82 percent to 91 percent for surveys and SSRS forms among grade 4-5 teachers. Differences between intervention and control group schools in baseline and year 1 survey response rates are not statistically significant at the 0.05 level. In year 2, however, teachers in intervention group schools were less likely to return surveys and SSRS forms than their counterparts in control group schools, with differences of 18–20 percentage points (72-73 percent compared with 91–93 percent). The difference in teacher survey response rates between intervention and control group schools is statistically significant at the 0.05 level.

Teacher implementation data—participating teachers in intervention schools

As described above (see note #12), the June implementation log data are critical for assessing the extent to which LIC lessons were delivered during the entire academic year, because log data collected in December or March cover a more limited time period. December and March implementation log data may still be useful for assessing the minimum number of LIC lessons implemented during the academic year. The accuracy of reported implementation of Daily Oral Language with Character and Writing with Character lessons depends on obtaining complete implementation logs at each collection point.

Among participating teachers in intervention schools, approximately 67 percent of teachers returned June implementation log data in year 1, and 81 percent of teachers did so in year 2 (see table 2.9). Approximately 84–85 percent of participating teachers returned some implementation log data. Mail, email, and telephone follow-ups were conducted to obtain additional information on core and supplementary program implementation. Considering all sources of lesson implementation data, 92–94 percent of participating teachers provided information on lesson implementation across year 1 and year 2.

Teacher implementation data—participating teachers in grades 4-5 in intervention schools

Implementation log return rates among participating teachers in grades 4-5 are similar to those among participating teachers in grades 2-5. Approximately 66 percent of grade 4-5 teachers returned June implementation log data in year 1, and 77 percent of teachers did so in year 2. Approximately 81–87 percent of participating grade 4-5 teachers returned some implementation log data and 91 percent of participating teachers provided some information on lesson implementation (any source) across year 1 and year 2.

Student-level data—primary analytic sample

Among students in the primary analytic sample, baseline SSRS forms were completed for 87 percent of eligible students,¹⁴ and year 2 SSRS forms were completed for 94 percent of students. The differences between intervention and control group schools in SSRS response rates were not statistically significant at the 0.05 level: the year 2 SSRS response rates were 88 percent in intervention group schools and 99 percent in control group schools.¹⁵ Although a difference could bias estimates of program impacts, such potential bias is likely reduced to some extent by the overall high response rate for the total sample (What Works Clearinghouse 2008). With an overall response rate of 94 percent and a difference of 11 percentage points between intervention and control groups, the potential bias in estimated impacts ranges from 0.05 standard deviation under optimistic assumptions to 0.10 under less optimistic assumptions, depending on the strength of the correlation between response and the outcome measure (What Works Clearinghouse 2008, p. 31). To further assess the potential for attrition-related bias, analyses of differences between intervention and control group schools in the retained primary analytic sample are described later in this chapter.

Year 2 survey data were collected for 96 percent of students in the primary analytic sample. Baseline test score data were available for 91 percent of students in the primary analytic sample, and year 2 test score data were available for 97 percent of students in the primary analytic sample. Differences in survey response rates and test score availability rates across intervention and control group schools were not statistically significant.

Student-level data—exploratory analytic sample

The bottom section of table 2.9 shows student SSRS and survey response rates and test score data availability rates for the exploratory analytic sample. Overall response rates range from 85 percent to 95 percent, with no differences between the intervention and control groups that are statistically significant at the 0.05 level.

¹⁴ For the purposes of calculating response rates, eligible students were defined as those who were enrolled in study schools when consent forms were distributed (the spring of year 0 and/or the fall of year 1) and were enrolled during the semester when data collection took place. Note that eligible students may not have been enrolled in a participating classroom at the time of data collection.

¹⁵ The difference between intervention-group and control-group schools in year 2 SSRS response rates was significant at the 0.054 confidence level.

Table 2.9. Response rates by data source and experimental group

| Measure and source | Overall | | Intervention | | Control | | Difference | z-test | p-value | | | |
|---|---------|---------|--------------|---------|---------|---------|---------------------|--------|---------|--|--|--|
| | Number | Percent | Number | Percent | Number | Percent | | | | | | |
| Teacher-level data—all eligible teachers^a | | | | | | | | | | | | |
| <i>Teacher survey^b</i> | | | | | | | | | | | | |
| Baseline | 680 | 90.29 | 325 | 89.85 | 355 | 90.70 | -0.75 | -0.17 | 0.86 | | | |
| Year 1 | 669 | 84.45 | 322 | 78.57 | 347 | 89.91 | -11.34 | -1.41 | 0.16 | | | |
| Year 2 | 625 | 81.60 | 299 | 71.24 | 326 | 91.10 | -19.86 | -1.70 | 0.09 | | | |
| <i>Implementation logs^{b,c}</i> | | | | | | | | | | | | |
| Year 1 any log | 296 | 84.12 | 296 | 84.12 | — | — | — | — | — | | | |
| Year 1 June log | 296 | 66.89 | 296 | 66.89 | — | — | — | — | — | | | |
| Year 1 any source | 296 | 92.23 | 296 | 92.23 | — | — | — | — | — | | | |
| Year 2 any log | 243 | 85.19 | 243 | 85.19 | — | — | — | — | — | | | |
| Year 2 June log | 243 | 81.07 | 243 | 81.07 | — | — | — | — | — | | | |
| Year 2 any source | 243 | 94.24 | 243 | 94.24 | — | — | — | — | — | | | |
| Teacher-level data—eligible 4th/5th grade teachers^d | | | | | | | | | | | | |
| <i>Teacher survey^e</i> | | | | | | | | | | | | |
| Baseline | 264 | 91.29 | 126 | 92.86 | 138 | 89.86 | 3.00 | 0.65 | 0.51 | | | |
| Year 1 | 258 | 81.78 | 124 | 76.61 | 134 | 86.57 | -9.96 | -1.07 | 0.28 | | | |
| Year 2 | 254 | 82.68 | 120 | 73.31 | 134 | 91.04 | -17.73 ^f | -2.09 | 0.04 | | | |
| <i>Any SSRS forms</i> | | | | | | | | | | | | |
| Baseline ^g | 137 | 91.24 | 65 | 92.31 | 72 | 90.28 | 2.03 | 0.47 | 0.64 | | | |
| Year 1 ^e | 258 | 81.78 | 124 | 77.42 | 134 | 85.82 | -8.40 | -0.80 | 0.42 | | | |
| Year 2 ^e | 260 | 82.69 | 126 | 72.22 | 134 | 92.54 | -20.32 | -1.89 | 0.06 | | | |
| <i>Implementation logs^{e,c}</i> | | | | | | | | | | | | |
| Year 1 any log | 116 | 87.07 | 116 | 87.07 | — | — | — | — | — | | | |
| Year 1 June log | 116 | 66.38 | 116 | 66.38 | — | — | — | — | — | | | |
| Year 1 any source | 116 | 91.38 | 116 | 91.38 | — | — | — | — | — | | | |
| Year 2 any log | 105 | 80.95 | 105 | 80.95 | — | — | — | — | — | | | |
| Year 2 June log | 105 | 77.14 | 105 | 77.14 | — | — | — | — | — | | | |
| Year 2 any source | 105 | 91.43 | 105 | 91.43 | — | — | — | — | — | | | |
| Student-level data—primary analytic sample^h | | | | | | | | | | | | |
| <i>Student SSRS</i> | | | | | | | | | | | | |
| Baseline ⁱ | 4,683 | 86.76 | 2,108 | 82.59 | 2,575 | 90.17 | -7.58 | -0.22 | 0.83 | | | |
| Year 2 ^e | 4,683 | 93.81 | 2,108 | 87.71 | 2,575 | 98.80 | -11.08 | -1.92 | 0.05 | | | |
| <i>Student survey</i> | | | | | | | | | | | | |
| Year 2—spring ^e | 4,683 | 96.41 | 2,108 | 95.45 | 2,575 | 97.20 | -1.76 | -1.01 | 0.31 | | | |
| <i>Test scores</i> | | | | | | | | | | | | |
| Year 0—spring ⁱ | 4,683 | 91.16 | 2,108 | 87.62 | 2,575 | 94.06 | -6.44 | -1.50 | 0.13 | | | |
| Year 2—spring ^e | 4,683 | 96.63 | 2,108 | 96.82 | 2,575 | 96.47 | 0.36 | 0.41 | 0.68 | | | |

| Measure and source | Overall | | Intervention | | Control | | Difference | z-test | p-value | | | |
|---|---------|---------|--------------|---------|---------|---------|------------|--------|---------|--|--|--|
| | Number | Percent | Number | Percent | Number | Percent | | | | | | |
| Student-level data—exploratory analytic sample^j | | | | | | | | | | | | |
| <i>Student SSRS</i> | | | | | | | | | | | | |
| Baseline ^k | 5,674 | 85.28 | 2,640 | 82.05 | 3,034 | 88.10 | -6.06 | -0.02 | 0.98 | | | |
| Year 1 ^e | 5,674 | 90.84 | 2,640 | 88.11 | 3,034 | 93.21 | -5.10 | -0.01 | 0.99 | | | |
| <i>Student survey</i> | | | | | | | | | | | | |
| Year 1—fall ^e | 5,674 | 88.40 | 2,640 | 86.78 | 3,034 | 89.82 | -3.04 | -0.06 | 0.95 | | | |
| Year 1—spring ^e | 5,674 | 95.45 | 2,640 | 95.00 | 3,034 | 95.85 | -0.85 | -0.51 | 0.61 | | | |
| <i>Test scores</i> | | | | | | | | | | | | |
| Year 0—spring ^k | 5,674 | 89.11 | 2,640 | 85.57 | 3,034 | 92.19 | -6.62 | -1.65 | 0.10 | | | |
| Year 1—spring ^e | 5,674 | 94.38 | 2,640 | 92.05 | 3,034 | 96.41 | -4.36 | -1.22 | 0.22 | | | |

Note: Test statistics (z-tests) are based on multilevel regression models that accounted for nesting of students within teachers (when applicable) and teachers within schools. No multiple comparison adjustments were applied.

a. Non-special education teachers in grades 2–5 in study schools (participants and nonparticipants).

b. Grades 2–5.

c. Calculated as the percentage of returned forms among participating teachers.

d. Non-special education teachers in grades 4–5 in study schools (participants and nonparticipants).

e. Grades 4–5.

f. Significantly different from 0 at the 0.05 level, two-tailed test.

g. Grade 4 only (SSRS data not collected from 5th grade teachers in year 0).

h. Grade 4 or 5 students in year 2 whose parents provided permission to participate and who were enrolled in participating classrooms in fall and spring of year 1 and spring of year 2.

i. Grades 2–3.

j. Grade 4 and 5 students whose parents provided permission to participate and who were enrolled in participating classrooms in fall and spring of year 1.

k. Grades 3–4.

Outcome domain measures

As described in chapter 1, both the primary confirmatory and exploratory research questions focus on potential program impacts on student academic achievement, social competence, and problem behaviors. For each domain, two to three outcome variables are analyzed. The student academic achievement measures analyzed are state English language arts standardized test scores and teacher reports of student academic competence. The social competence measures include teacher reports on the SSRS (total scores) and student reports of altruism and empathy. Teacher reports of students' externalizing behavior and student reports of aggression and delinquent behavior are analyzed for the problem behaviors domain.

Sample characteristics

Schools

Table 2.10 presents characteristics of the schools that were randomly assigned to the experimental group. Of the 50 participating schools, 34 were in the Los Angeles and San Diego Metropolitan Statistical Areas, 13 in northern California, and 3 in central California; 30 schools were in a large city or urban fringe of a large city, 6 in a midsize city (population of 25,000–250,000), and 14 in a small town or rural area (not shown). Overall, the average Academic Performance Index of the schools in the sample was 778, which was very close to the state average of 776 (California Department of Education 2010b). Approximately 59 percent of the students served by the schools were eligible for free or reduced-price meals, 49 percent were

classified as Hispanic, 32 percent were classified as non-Hispanic White, and 31 percent were classified as English language learner students. Enrollments in schools included in the study averaged 354 students. As shown in table 2.10, no statistically significant differences in school characteristics were found between schools in the intervention group and schools in the control group. The results are unchanged when examining the sample of 49 schools that were retained in the study during the full implementation period.

Table 2.10. School-level baseline characteristics of study schools (randomized sample)

| Characteristic | Baseline means (standard deviation) | | | Difference | t-test ^a | p-value |
|-------------------------------------|--|--------------------|--------------------|------------|---------------------|---------|
| | Overall | Intervention group | Control group | | | |
| Academic Performance Index | 778.20 (76.11) | 780.68 (88.10) | 775.72 (63.66) | 4.96 | 0.23 | 0.82 |
| African American (percent) | 6.94 (13.11) | 9.16 (17.48) | 4.72 (5.93) | 4.44 | 1.20 | 0.23 |
| Asian (percent) | 5.78 (10.91) | 6.40 (12.56) | 5.16 (9.19) | 1.24 | 0.40 | 0.69 |
| Hispanic (percent) | 49.34 (28.57) | 47.88 (28.82) | 50.80 (28.84) | -2.92 | -0.36 | 0.72 |
| Non-Hispanic White (percent) | 32.50 (29.76) | 30.72 (29.31) | 34.28 (30.70) | -3.56 | -0.42 | 0.68 |
| Free/reduced-price meals (percent) | 58.64 (27.83) | 59.04 (29.09) | 58.24 (27.10) | 0.80 | 0.10 | 0.92 |
| English language learners (percent) | 30.78 (22.98) | 30.24 (22.66) | 31.24 (23.74) | -1.08 | -0.16 | 0.87 |
| Parent education ^b | 2.68 (0.70) | 2.67 (0.71) | 2.70 (0.70) | -0.04 | -0.19 | 0.85 |
| Enrollment | 353.52 (154.85) | 335.52 (143.16) | 371.12 (166.78) | -35.20 | -0.80 | 0.43 |
| Sample size (schools) | 50 | 25 | 25 | | | |

a. An independent samples *t*-test was used to estimate differences in baseline means between intervention and control groups. No multiple comparison adjustments were applied. No differences were statistically significant at the 0.05 level (two-tail).

b. Average education level of the parents of students in the school based on an ordinal variable ranging from 1 to 5 (1 = less than high school, 2 = high school graduate, 3 = some college, 4 = college graduate, 5 = graduate school).

Source: California Department of Education 2009; U.S. Department of Education 2009.

Teachers

As described above, general education teachers of grades 2–5 in elementary schools were targeted for study participation. Overall, 87–95 percent of grade-eligible teachers in study schools participated in the study over the three-year data collection period (see table 2.3). Tables 2.11 and 2.12 present characteristics of teachers who participated in the study. Because class enrollment sizes are typically smaller in grades 2 and 3 than in grades 4 and 5 in California, the teacher sample comprises proportionately more grade 2 and 3 teachers (61 percent) than grade 4 and 5 teachers (39 percent). Approximately 85 percent of participating teachers were female, and 85 percent reported that they were non-Hispanic White. The intervention and control groups did not differ on grade, gender, or ethnic composition in the randomized sample, the year 1 sample (table E1 in appendix E), or the year 2 sample (table E2 in appendix E). Teachers participating in the study had an average of 13 years of total teaching experience and 8 years of experience teaching in the school that they were working in at the time of the study. None of the differences

across intervention and control schools in teacher baseline measures presented in tables 2.11 and 2.12 were statistically significant.

Table 2.11. Teacher baseline demographic characteristics by experimental group (teacher participants in baseline year)

| Characteristic | Baseline percentages | | | | | |
|-----------------------------------|----------------------|--------------------|---------------|------------|-----------------------------|-----------------|
| | Overall | Intervention group | Control group | Difference | <i>z</i> -test ^a | <i>p</i> -value |
| <i>Grade taught (year 0)</i> | | | | | | |
| 2 nd | 33.33 | 33.97 | 32.74 | 1.23 | 0.03 | 0.98 |
| 3 rd | 27.93 | 26.92 | 28.87 | -1.95 | | |
| 4 th | 20.22 | 20.19 | 20.24 | -0.05 | | |
| 5 th | 18.52 | 18.91 | 18.15 | 0.76 | | |
| <i>Gender</i> | | | | | | |
| Female | 84.57 | 85.90 | 83.33 | 2.57 | 0.71 | 0.48 |
| Male | 15.43 | 14.10 | 16.67 | -2.57 | | |
| <i>Race/ethnicity^b</i> | | | | | | |
| Asian | 6.03 | 7.19 | 4.97 | 2.22 | 1.03 | 0.30 |
| African American | 6.03 | 7.53 | 4.66 | 2.87 | 1.09 | 0.27 |
| Hispanic | 16.93 | 16.55 | 17.27 | -0.72 | -0.13 | 0.89 |
| White | 84.78 | 81.69 | 87.54 | -5.85 | -1.37 | 0.17 |
| <i>Sample size</i> | | | | | | |
| Schools | 50 | 25 | 25 | | | |
| Teachers ^c | 648 | 312 | 336 | | | |

a. A two-tailed test that adjusted for intraclass correlations was used to test differences in baseline characteristics across intervention and control group schools. No multiple comparison adjustments were applied. No differences were statistically significant at the 0.05 level.

b. Racial/ethnic groups are not mutually exclusive.

c. Complete data were available for teacher grade and gender. Approximately 3.7 percent of teacher responses were missing on the race/ethnicity items.

Source: Baseline teacher survey and teacher rosters.

Table 2.12. Teacher baseline measures by experimental group (teacher participants)

| Measure | Baseline means (standard deviation) | | | Difference | z-test ^a | p-value |
|-------------------------------------|--|--------------------|-----------------|------------|---------------------|---------|
| | Overall | Intervention group | Control group | | | |
| <i>Randomized sample</i> | | | | | | |
| Years of teaching experience | 12.74 (8.95) | 12.24 (9.14) | 13.18 (8.77) | -0.94 | -0.80 | 0.42 |
| Years of teaching in school | 7.84 (6.72) | 7.18 (6.39) | 8.43 (6.96) | -1.25 | -1.25 | 0.21 |
| Student feelings of belonging | 3.50 (0.67) | 3.48 (0.69) | 3.53 (0.65) | -0.05 | -0.47 | 0.64 |
| School expectations | 4.44 (0.55) | 4.46 (0.53) | 4.42 (0.57) | 0.04 | 0.45 | 0.65 |
| Sample size (schools) | 50 | 25 | 25 | | | |
| Sample size ^b (teachers) | 614 | 292 | 322 | | | |
| <i>Year 1 sample</i> | | | | | | |
| Years of teaching experience | 12.69 (8.81) | 12.11 (8.92) | 13.24 (8.69) | -1.13 | -1.02 | 0.31 |
| Years of teaching in school | 7.69 (6.62) | 6.91 (6.10) | 8.43 (7.00) | -1.52 | -1.58 | 0.12 |
| Student feelings of belonging | 3.52 (0.64) | 3.52 (0.64) | 3.53 (0.65) | -0.01 | 0.09 | 0.93 |
| School expectations | 4.46 (0.54) | 4.48 (0.52) | 4.45 (0.53) | 0.03 | 0.26 | 0.80 |
| Sample size (schools) | 50 | 25 | 25 | | | |
| Sample size ^b (teachers) | 557 | 269 | 288 | | | |
| <i>Year 2 sample</i> | | | | | | |
| Years of teaching experience | 12.78 (8.57) | 12.35 (8.49) | 13.14 (8.64) | -0.79 | -0.71 | 0.48 |
| Years of teaching in school | 8.03 (6.63) | 7.20 (5.97) | 8.72 (7.06) | -1.52 | -1.42 | 0.15 |
| Student feelings of belonging | 3.56 (0.64) | 3.55 (0.64) | 3.57 (0.64) | -0.02 | -0.10 | 0.92 |
| School expectations | 4.45 (0.52) | 4.44 (0.53) | 4.46 (0.51) | -0.02 | -0.52 | 0.60 |
| Sample size ^b (teachers) | 464 | 209 | 255 | | | |

a. A two-tailed test that adjusted for intraclass correlations was used to test differences in baseline characteristics across intervention and control group schools. No multiple comparison adjustments were applied.

b. Teacher-level missing data rates range from 3.8 percent to 5.9 percent.

Source: Baseline teacher survey and teacher rosters.

Students

Student study participants were equally represented in each grade and gender (table 2.13). Approximately 55 percent of participating students were of Hispanic origin, 27 percent were non-Hispanic White, 8-9 percent were Asian, and 6-7 percent were Black. Some 37-38 percent of students were classified as English language learners, and 9-11 percent were classified as special education students. Approximately 17 percent of students had parents who did not have a high school degree, 23 percent had parents with a high school degree, 21-22 percent had parents who attended college but did not graduate, 24-25 percent had parents who graduated from college, and 14 percent had missing parent education information. The intervention and control groups did not differ by school grade, gender, race/ethnicity, English language learner status,

parent education, or special education status. No statistically significant differences were detected between the intervention and control groups for the primary analytic and exploratory analytic samples as well.

Table 2.14 shows means and standard deviations of baseline measures of outcome variables overall and separately for all consenting students in intervention and control group schools for the retained primary analytic and exploratory analytic samples. Across both samples, participating students in intervention and control group schools reported similar levels on all of the baseline outcome variables.

Table 2.13. Student demographic characteristics by experimental group

| Characteristic | Baseline percentages | | | Difference | Test statistic ^a (df) | p-value | | | |
|--|----------------------|--------------------|---------------|------------|-------------------------------------|---------|--|--|--|
| | Overall | Intervention group | Control group | | | | | | |
| Primary analytic sample | | | | | | | | | |
| <i>School grade (year 2)</i> | | | | | | | | | |
| 4 | 49.67 | 49.05 | 50.17 | -1.12 | 0.19 | 0.85 | | | |
| 5 | 50.33 | 50.95 | 49.83 | 1.12 | | | | | |
| <i>Gender</i> | | | | | | | | | |
| Female | 50.22 | 50.00 | 50.41 | -0.41 | -0.18 | 0.86 | | | |
| Male | 49.78 | 50.00 | 49.59 | 0.41 | | | | | |
| <i>Race/ethnicity</i> | | | | | | | | | |
| Asian | 8.78 | 9.20 | 8.43 | 0.77 | 0.97 _(5,47) | 0.44 | | | |
| Black | 6.32 | 9.96 | 3.34 | 6.62 | | | | | |
| Hispanic | 55.80 | 52.09 | 58.83 | -6.74 | | | | | |
| White | 26.76 | 25.95 | 27.42 | -1.47 | | | | | |
| Other | 2.33 | 2.75 | 1.98 | 0.77 | | | | | |
| Missing | 0.02 | 0.05 | 0.00 | 0.05 | | | | | |
| <i>English language learner status</i> | | | | | | | | | |
| English language learner | 38.44 | 36.95 | 39.65 | -2.70 | 0.09 | 0.93 | | | |
| <i>Parent education</i> | | | | | | | | | |
| No high school degree | 16.95 | 16.08 | 17.67 | -1.59 | -0.06 | 0.95 | | | |
| High school degree | 23.06 | 23.91 | 22.37 | 1.54 | | | | | |
| Some college | 21.33 | 21.82 | 20.93 | 0.89 | | | | | |
| College degree or more | 24.73 | 23.96 | 25.36 | -1.40 | | | | | |
| Missing | 13.92 | 14.23 | 13.67 | 0.56 | | | | | |
| <i>Special education status</i> | | | | | | | | | |
| Special education | 10.57 | 10.11 | 10.95 | -0.84 | -0.03 | 0.98 | | | |
| <i>Sample size</i> | | | | | | | | | |
| Students ^b | 4,683 | 2,108 | 2,575 | | | | | | |
| Exploratory analytic sample | | | | | | | | | |
| <i>School grade (year 1)</i> | | | | | | | | | |
| 4 | 50.90 | 51.48 | 50.40 | 1.08 | -0.10 | 0.92 | | | |
| 5 | 49.10 | 48.52 | 49.60 | -1.08 | | | | | |
| <i>Gender</i> | | | | | | | | | |
| Female | 49.74 | 49.55 | 49.90 | -0.35 | 0.14 | 0.89 | | | |
| Male | 50.26 | 50.45 | 50.10 | 0.35 | | | | | |
| <i>Race/ethnicity</i> | | | | | | | | | |
| Asian | 8.18 | 8.45 | 7.94 | 0.51 | 1.13 _(5, 48) | 0.36 | | | |
| Black | 6.96 | 11.14 | 3.33 | 7.81 | | | | | |
| Hispanic | 54.81 | 49.77 | 59.20 | -9.43 | | | | | |
| White | 27.12 | 26.78 | 27.42 | -0.64 | | | | | |
| Other | 2.29 | 2.54 | 2.08 | 0.46 | | | | | |

| Characteristic | Baseline percentages | | | | Test statistic^a (df) | p-value |
|--|-----------------------------|---------------------------|----------------------|-------------------|--|----------------|
| | Overall | Intervention group | Control group | Difference | | |
| Missing | 0.63 | 1.33 | 0.03 | 1.30 | | |
| <i>English language learner status</i> | | | | | | |
| English language learner | 36.54 | 33.64 | 39.06 | -5.42 | -0.44 | 0.66 |
| <i>Parent education</i> | | | | | | |
| No high school degree | 17.50 | 15.68 | 19.08 | -3.40 | 0.79 | 0.43 |
| High school degree | 23.02 | 23.41 | 22.68 | 0.73 | | |
| Some college | 22.45 | 21.52 | 23.27 | -1.75 | | |
| College degree or more | 23.51 | 24.63 | 22.54 | 2.09 | | |
| Missing | 13.52 | 14.77 | 12.43 | 2.34 | | |
| <i>Special education status</i> | | | | | | |
| Special education | 8.97 | 9.40 | 8.61 | 0.79 | 0.81 | 0.42 |
| <i>Sample size</i> | | | | | | |
| Schools | 50 | 25 | 25 | | | |
| Students ^c | 5,674 | 2,640 | 3,034 | | | |

a. For all characteristics except student race/ethnicity, a two-tailed test that adjusted for intraclass correlations was used to test differences in baseline characteristics across intervention and control schools (*z*-test). To test for differences in racial/ethnic composition, school-level proportions for all six of the racial/ethnic categories were calculated. Intervention/control school differences in racial/ethnic composition were estimated using a multivariate regression model and an F-test (degrees of freedom in parentheses) was used to test for group differences. No multiple comparison adjustments were applied.

b. One student was missing data for the special education variable.

c. Forty-three students were missing data for the special education variable.

Source: Student roster data and district archival records.

Table 2.14. Student baseline measures by experimental group

| Characteristic | Baseline means (standard deviation) | | | Difference | z-test ^a | p-value |
|------------------------------------|--|--------------------|--------------------|------------|---------------------|---------|
| | Overall | Intervention group | Control group | | | |
| Primary analytic sample | | | | | | |
| English language arts test | 0.050 (0.923) | 0.065 (0.928) | 0.039 (0.919) | 0.026 | 0.28 | 0.78 |
| Academic competence | 31.423 (9.237) | 31.735 (9.265) | 31.187 (9.211) | 0.547 | 1.25 | 0.21 |
| Social skills | 43.229 (12.311) | 42.988 (12.128) | 43.410 (12.446) | -0.422 | -0.12 | 0.90 |
| Altruism | 4.384 (0.605) | 4.346 (0.602) | 4.416 (0.606) | -0.070 | -1.56 | 0.12 |
| Empathy | 2.356 (0.345) | 2.354 (0.339) | 2.358 (0.350) | -0.004 | 0.12 | 0.90 |
| Externalizing | 2.081 (2.798) | 2.214 (2.885) | 1.981 (2.728) | 0.232 | 1.20 | 0.23 |
| Aggression | 1.175 (0.369) | 1.193 (0.375) | 1.161 (0.363) | 0.032 | 0.73 | 0.46 |
| Delinquent behavior | 1.105 (0.253) | 1.112 (0.253) | 1.100 (0.253) | 0.012 | 0.56 | 0.58 |
| School expectations | 3.981 (0.783) | 3.940 (0.795) | 4.013 (0.772) | -0.073 | -0.72 | 0.47 |
| Students' feelings of belonging | 3.596 (0.655) | 3.556 (0.639) | 3.628 (0.666) | -0.073 | -1.00 | 0.32 |
| <i>Sample size</i> | | | | | | |
| Students ^b | 4,683 | 2,108 | 2,575 | | | |
| Exploratory analytic sample | | | | | | |
| English language arts test | 0.037 (0.929) | 0.097 (0.946) | -0.012 (0.912) | 0.109 | 0.62 | 0.54 |
| Academic competence | 31.102 (9.643) | 31.244 (9.766) | 30.986 (9.541) | 0.258 | 0.63 | 0.53 |
| Social skills | 43.406 (13.254) | 43.298 (13.805) | 43.494 (12.792) | -0.195 | -0.45 | 0.66 |
| Altruism | 4.327 (0.621) | 4.311 (0.622) | 4.340 (0.619) | -0.029 | -0.61 | 0.54 |
| Empathy | 2.327 (0.354) | 2.326 (0.351) | 2.329 (0.356) | -0.003 | 0.23 | 0.82 |
| Externalizing | 2.157 (2.898) | 2.253 (3.001) | 2.078 (2.810) | 0.175 | 1.08 | 0.28 |
| Aggression | 1.215 (0.410) | 1.214 (0.406) | 1.215 (0.413) | -0.001 | -0.39 | 0.70 |
| Delinquent behavior | 1.120 (0.268) | 1.114 (0.259) | 1.124 (0.275) | -0.010 | -0.81 | 0.42 |
| School expectations | 3.975 (0.799) | 3.960 (0.810) | 3.988 (0.790) | -0.028 | -0.14 | 0.89 |
| Students' feelings of belonging | 3.513 (0.668) | 3.490 (0.650) | 3.531 (0.683) | -0.041 | -0.27 | 0.79 |
| <i>Sample size</i> | | | | | | |
| Schools | 50 | 25 | 25 | | | |
| Students ^c | 5,674 | 2,640 | 3,034 | | | |

a. A two-tailed test that adjusted for intraclass correlations was used to test differences in baseline characteristics across intervention and control schools. No multiple comparison adjustments were applied.

b. Missing data rates range from 8.0 percent to 13.8 percent.

c. Missing data rates range from 10.0 percent to 15.3 percent.

Source: Baseline district test score data, teacher SSRS surveys, and student surveys.

Data analysis methods

Impacts of the LIC program were estimated by comparing post-intervention outcomes for students and classrooms in the intervention group to the outcomes for their counterparts in the control group. The primary hypothesis-testing involved fitting conditional multilevel regression models (that is, hierarchical linear modeling), with additional terms to account for the nesting of individuals within higher units of aggregation (see, for example, Goldstein 1987; Raudenbush and Bryk, 2002; Murray 1998). The study involved school-level random assignment and delivery of training courses to teachers within intervention group schools, who in turn had the opportunity to incorporate 25 supplementary lessons into their classroom instruction during the academic year. Random effects of school site and teacher were included in the models to account for the nesting of observations within schools and teachers, respectively. Fixed effects include intervention group, baseline (pretest) measures of outcome variables, randomization strata, and other individual and aggregate school-level covariates. The purpose of including statistical controls is to minimize random error and to increase the precision of the estimates.

As an illustrative example, the following type of three-level hierarchical linear modeling for a continuous outcome was estimated:

$$Character_{ijkl} = \alpha_0 + \beta_1 Tx_{kl} + \sum \beta_I I_{ijkl} + \sum \beta_S S_{kl} + \sum_{l=2}^L \beta_{St} Strata_l + \mu_{jkl} + \nu_{kl} + \varepsilon_{ijkl} \quad (1)$$

where subscripts i , j , k , and l denote student, teacher, school, and strata, respectively; $Character$ represents the student outcome variable; Tx is a dichotomous variable indicating student attendance at the school in the intervention condition; I represents a vector of student-level control variables measured prior to random assignment (including a baseline measure of the outcome variable or measure closely related to the outcome variable); S is a set of school-level control variables (including an aggregated, school-level baseline measure of the outcome variable); and $Strata$ is a set of dichotomous variables representing fixed effects for strata. Lastly, ε_{ijkl} , μ_{jkl} , and ν_{kl} are error terms for students, teachers, and schools, respectively. In this model, the intervention effect is represented by β_1 , which captures adjusted intervention-control group school differences in the outcome variable. With one intervention impact coefficient, larger strata/schools are given more weight than smaller strata/schools in the estimation of the aggregate intervention effect. The model includes two baseline measures of the outcome variable as covariates, both assessed prior to random assignment: an aggregated, school-level baseline measure and a student-level measure.¹⁶ Including both covariates enhances precision of the impact estimate.¹⁷

¹⁶ When analyzing student standardized test scores, both student- and school-level standardized test scores from the spring prior to the first implementation year were included as covariates in the impact analysis model. Because standardized test scores are not vertically aligned across grades, these covariates should not be conceived of as “pretest measures of the outcome variable” because they measure different academic skills than those assessed after program implementation. Moreover, because standardized test scores are not comparable across grades, estimated impacts should not be interpreted as reflecting intervention-control group differences in growth in performance over

As mentioned above, the analytic sample for the primary confirmatory analyses was restricted to students in grades 4 and 5 in year 2 so that outcomes based on self-reported student data could be included. The study sample was also restricted to students exposed for two consecutive years to teachers participating in the study. Thus, two-year LIC exposure impacts are estimated.

Specific pre-intervention covariates were selected based on study design considerations and judgments about the extent to which they would explain variance in the outcome variables. The following pre-intervention covariates were included in the models examining primary student outcomes:¹⁸

Student-level

- Pretest measure of each outcome variable¹⁹
- Grade in school (4 or 5)
- Gender
- Race/ethnicity (non-Hispanic White, Asian, Black, Hispanic, Other)
- English language learner status
- Parent education (no high school degree, high school degree, some college, college degree or more)

School-level

- Mean pretest measure of outcome variable among grade 2 and 3 students at the school
- School academic performance (California's Academic Performance Index)
- Dichotomous variables for each stratum (one dichotomous variable omitted)

To examine the intermediate teacher-reported measures used to address research questions 4 and 5, multilevel regression models analogous to equation 1 were also used to assess teacher-level measures of students' feelings of belonging and school expectations. For example, for teacher reports of students' feelings of belonging, the following model was estimated:

$$Belonging_{jkl} = \alpha_0 + \beta_1 T_{x_{kl}} + \sum \beta_T T_{jkl} + \sum \beta_S S_{kl} + \sum_{l=2}^l \beta_{ST} Strata_l + \nu_{kl} + \mu_{jkl} \quad (2)$$

where the subscripts and variables are defined the same way as they are for equation 1, except that *Belonging* represents the teacher outcome variable and *T* represents a vector of teacher-level control variables measured prior to random assignment (including pre-intervention measures of the outcome variable). The teacher-level models included the following covariates:²⁰

time (Singer and Willett 2003). Instead, such impacts reflect covariate-adjusted intervention-control group differences in post-implementation means.

¹⁷ Student-level pretest measures based on the student surveys were not included as covariates in the model because such measures were unavailable for students who were in grade 3 during the fall of the first implementation year.

¹⁸ Appendix F provides details concerning the covariates included in specific models.

¹⁹ The applicable mean of pretest measure of grade 4 and 5 students at the school is subtracted from each student-level pretest measure.

²⁰ See appendix F for more details.

Teacher-level

- Pretest measure of teacher reports of students' feelings of belonging and school expectations²¹
- Teacher-aggregated measure of student reports of feelings of belonging and school expectations
- Grade taught
- Years employed as a teacher at the school
- Gender
- Race/ethnicity (non-Hispanic White, Asian, Black, Hispanic, other)

School-level

- Mean pretest measure of outcome variable in the school (based on grade 4 and 5 teachers only)
- School academic performance (California's Academic Performance Index)
- Dichotomous variables for each stratum

To examine the intermediate student-reported measures of belonging and student expectations, a three-level hierarchical linear model analogous to equation 1 was estimated, where students are nested within teachers and teachers are nested within schools.

Analogous student-and teacher-level models were estimated to examine the exploratory research questions. To investigate the impacts of exposure to the LIC program during the first year of program implementation, spring of year 1 outcomes were examined and the analytic sample was restricted to participants in study schools in the spring of year 1. All estimates were obtained using the Stata 10.1 statistical package (StataCorp 2007). To obtain impact estimates, the multilevel mixed effects restricted maximum likelihood estimator implemented in Stata was used.²²

Sensitivity analyses

Several types of sensitivity analyses were conducted to test the robustness of the results across different models and samples. Appendixes G, H, I, and J present the sensitivity analyses. The following types of sensitivity analyses were conducted:

Impact models with fewer baseline covariates. Two sets of models were estimated that included fewer covariates than the benchmark model described above—one model that included randomization strata and intervention status as covariates and a second model that included randomization strata, intervention status, and a pretest measure of the outcome variable as a covariate.

Impact models using different missing data-handling methods. As described in the following section, multiple imputation techniques were used to impute missing covariate values for the benchmark model. To examine the possible consequences of using different missing data-handling techniques for the impact estimates, estimates were obtained from four different

²¹ The applicable mean pretest in the school is subtracted from the teacher-level pretest.

²² Unlike some other software packages, Stata's multilevel modeling commands calculate *z*-tests instead of *t*-tests for hypothesis testing because the finite *t*-distribution does not have a simple form in variance components models (Rabe-Hesketh and Skrondal 2008).

models/analytic samples: (1) multilevel analysis of covariance impact estimates based on cases with complete data on baseline and post-implementation measures (that is, listwise deletion); (2) multilevel analysis of covariance impact estimates based on cases with complete data on post-implementation measures only; (3) multilevel analysis of covariance impact estimates based on the same sample as (2), but with baseline measures and missing-value dummy variables included in the model; and (4) multilevel analysis of covariance results in which both the baseline measures and the outcome variable are imputed.

Impact models based on different analytic samples. The analytic samples that are the basis of the benchmark model included students who migrated across study schools. As noted above, some of these students crossed over from control to intervention schools (and vice versa) during the study period (crossovers), while others moved to schools that were in the same study condition as the school that they migrated from (non-crossovers). For the benchmark models, crossover students were analyzed as though they were in their originally assigned condition. Models that excluded crossover students and models that excluded students that migrated across study school (both crossovers and non-crossovers) were estimated as a sensitivity check.

Impact models based on samples that exclude nonintact strata. In an attempt to ascertain how school-level attrition may have biased impact estimates from the benchmark model, models that excluded the stratum in which a school dropped out of the study in the analytic sample were estimated. Estimates based on an analytic sample that excludes nonintact randomization strata are unlikely to be biased due to school-level attrition but are likely to exhibit more random error due to the reduced sample size. Statistical power estimates for such models are provided in the discussion of sampling and power in the next section.

Impact models based on weighted data. As described above, a non-response subsampling strategy was used to follow-up with securing consent forms from cohort 1 parents who had not yet returned consent forms. Up to five parents among all grade 2-5 students who did not return consent forms were randomly sampled within each classroom for more intensive follow-up. To ascertain the extent to which nonreturn of consent forms among cohort 1 students may have biased estimates of impacts, sensitivity analyses were conducted that applied weights to successfully recruited students in the subsample so that such students would represent those who did not return informed consent forms but were not randomly sampled. In these analyses, successfully recruited students were weighted by the inverse of the sampling rate in each class.²³ Impact analyses based on weighted data are presented in appendixes H and J.

The nonresponse subsampling strategy used for cohort 1 schools was as follows: 50 percent of eligible students were randomly sampled when there were 10 or fewer outstanding consent forms in a classroom,²⁴ and 5 students were sampled when there were more than 10 outstanding consent forms in a classroom. This resulted in a sampling rate of 48 percent among students or

²³ For example, if 11 students in a classroom did not return consent forms, 5 of those 11 students were sampled for more intensive follow-up. If three of the five sampled students were successfully recruited, the three cases were weighted by the inverse of the within-class sampling rate, or 11/5. Students who had already returned consent forms received a weight of 1.

²⁴ Values were rounded up when determining the number of students with whom to follow up more intensively. For example, if there were seven outstanding consent forms, four students were randomly sampled for more intensive follow-up.

parents who did not return consent forms. Of the parents who did not initially return consent forms to allow their children to participate, 61 percent returned consent forms following these more intensive recruitment efforts. Among sampled parents who returned consent forms, 80 percent provided affirmative consent. The implementation of the subsampling strategy resulted in a weighted consent return rate of 80 percent and a weighted student participation rate of 66 percent at the end of the first implementation year, compared with 60 percent and 50 percent at the beginning of year 1. In total, 355 additional grade 2-5 students participated in the study as a consequence of these follow-up efforts.

Multiple hypothesis testing procedures

The procedures described by Schochet (2008) were used to account for multiple hypothesis tests involving the numerous outcome variables assessed in the study. Specifically, three primary student outcome domains were delineated to reflect key outcomes—academic performance, social competence, and problem behaviors. Within each of these domains, two to three outcomes were examined. Academic achievement measures included state English language arts standardized test scores and student academic competence; social competence included the SSRS (total scores), altruism, and empathy; and problem behaviors included externalizing behavior, aggression, and delinquent behavior. Within each outcome domain, multiple comparison procedures were used in the confirmatory impact analyses to adjust for the inflation of type 1 errors due to hypothesis tests across the tests. Thus, adjustments were applied to adjust for the two statistical tests for the academic performance outcomes and for the three tests for each of the social competence and problem behavior outcomes. Statistical tests within each primary student outcome domain (e.g., academic performance) did not adjust for the tests with the other student outcome domain (e.g., problem behaviors).

Similarly, four intermediate classroom-level outcomes were analyzed to establish part of the theory of action linking the intervention to student academic achievement, social competence, and problem behaviors: teacher reports of students' feelings of belonging, student reports of feelings of belonging, teacher reports of school expectations, and student reports of school expectations. The two intermediate outcomes based on teacher reports were examined for two samples: the overall sample teachers in grades 2-5 and the subsample of teachers in grades 4-5. Multiple comparison procedures were used to adjust for the inflation of type 1 errors across these six statistical tests.

If statistically significant impacts were detected prior to adjusting for multiple testing, bootstrap resampling procedures were to be used to adjust for multiple testing within the primary and intermediate outcome domains (Westfall and Young 1993). Bootstrap resampling methods for multiple hypothesis tests provide strong control of the familywise error rate and account for the correlations between tests, which can reduce some of the adverse effect of multiple testing on statistical power (Schochet 2008). If (unadjusted) statistically significant impacts were not detected, the *p*-values were adjusted using the Bonferroni adjustment procedure.

Finally, exploratory analyses were conducted to estimate impacts of exposure to the LIC program during the first year of program implementation. Because these analyses were exploratory and not used to make inferences about program impacts, no adjustments for multiple hypotheses were made in the exploratory analyses.

Missing data

The primary analyses relied on multiple imputation techniques to deal with covariates with missing values. Parameter estimates and standard errors from multiple imputation are reported in chapters 4 and 5 of this report. As described above, sensitivity analyses were conducted to examine how robust the findings were with respect to missing data-handling procedures.

The primary and intermediate confirmatory analyses used Schafer's (1997) multiple imputation techniques. Schafer's approach involves imputing missing values by using a prediction model that includes predictors and response variables from the substantive model as well as other variables. To incorporate random variation in the imputations of variables in the substantive model, a Bayesian method (data augmentation) is used to iterate between random imputations under a specified set of parameter values and random draws from the posterior distribution of the parameters. This entire process is done several times to produce several complete data sets. Substantive analyses are done on each of the complete data sets, parameter values are averaged, and standard errors are calculated.

One complication with this approach is that it does not appear to be straightforward to account for the clustering of students within classrooms and schools. The workaround used for this study was to include dummy variables for the $N-1$ schools in the imputation model.

For the primary analyses, only pre-intervention measures were imputed. The imputation model included all the variables in the impact analysis model, including each primary outcome variable, dummy variables for schools, and other variables related to the missing data process. Outcome variables were used in the imputation model but were not imputed (von Hippel 2007). All the outcomes across all the analytic samples were included in the imputation model. To preserve interactions between intervention status and other variables in the impact analysis models, the dataset was split into intervention and control group samples, and separate imputation models were run for each group.

3. Implementation of Lessons in Character intervention

This chapter presents a description of the Lessons in Character (LIC) intervention. It describes the intended implementation of the intervention for the impact study. It then describes how the program was implemented in intervention group schools, including details about participation in training sessions, delivery of core and supplementary lessons in classrooms, and the presence of observable intervention artifacts in classrooms and public spaces. The chapter concludes with a description of reported character development activities in intervention and control group schools.

Intended Lessons in Character program implementation

Training and coaching

Teachers in schools randomly selected to be in the intervention-group condition had the opportunity to participate in a one-day training in the Lessons in Character program and to receive about two hours of coaching support for delivery of the curricular material during the fall semester in year 1 of the study. Dr. B. David Brooks, the developer of LIC, conducted the training and onsite coaching sessions. After the training, Dr. Brooks moderated an email discussion group for all teachers in the intervention group and provided monthly tips for implementation.

The framework for the one-day training was based on the Partnership for Character Education's "Eleven Principles of Effective Character Education," which outlines 11 basic principles that suggest how character will look when implemented in a classroom or school. Throughout the training, character education was defined as the systematic, purposeful teaching of core consensus values that lead to habits of good character. The training was designed to familiarize teachers with the rationale for incorporating character education in a systematic way into their daily schedule through the use of literature and lessons that allow students to learn about what being a person of character is and by practicing the skills learned in the lessons. The training emphasized that teaching about character was not another job but rather reinforcement that they already teach students the habits of good character, and that they should be mindful of the influence teachers have on their students as a model of good character. The training included an overview of character education and a hands-on review of the LIC, Daily Oral Language with Character, and Writing with Character materials. The STAR process was reviewed, and ways to incorporate it into the language and culture of the school and infuse it into the curriculum were explored. Participants were then divided into grade-specific breakout groups to work on activities from the materials for their grade level. The training ended with school teams working together to outline plans for implementing LIC, Daily Oral Language with Character, and Writing with Character at their respective sites.

The onsite coaching provided by Dr. Brooks was designed to be informal, involving observations of an LIC lesson in each school and meetings to discuss teachers' experiences, questions, comments, and challenges in integrating LIC into the regular teaching day. Dr. Brooks visited all 25 intervention group schools in the late fall and early winter of the first implementation year. Coaching was delivered either individually to teachers or in groups.

According to the developer, delivery of the LIC lesson materials does not require extensive teacher preparation or training. The training and coaching support provided to teachers participating in the evaluation study was likely more intensive than that typically received by teachers who implement LIC.

Core and supplementary curricular material

As described above, the LIC core curriculum consists of 25 20-minute lessons designed to be delivered weekly during the course of the academic year, with one lesson focusing on the STAR decision-making model and four lessons focusing on each of the six character themes.

Participating teachers in intervention group schools were asked to complete, at a minimum, the STAR lesson and the first three lessons for each of the six themes, for a total of 19 lessons per academic year.²⁵ The supplementary lessons include 180 5-minute Daily Oral Language with Character lessons and 36 20-minute Writing with Character lessons (grades 3–5 only). Although all grade-eligible intervention group teachers received the supplementary materials, both Daily Oral Language with Character and Writing with Character lessons were implemented voluntarily.

Schoolwide Lessons in Character activities

As described above, administrators of participating schools received an implementation guide that provides actionable guidance on how to infuse character development strategies into the school discipline policy, how to create a school climate based on good character, and how to involve families and the community in character education efforts. Moreover, participating teachers and school administrators received posters displaying the STAR model and the six character themes. Thus, all intervention group schools had the opportunity to implement schoolwide character education–related activities and display LIC artifacts in classrooms and public spaces to reinforce the character themes and the decision-making model.

Actual Lessons in Character program implementation

Training

During the late summer prior to program implementation, 19 one-day professional development sessions involving the 25 intervention group schools were held in locations close to participating schools. Table 3.1 shows the percentages of eligible teachers and principals who participated in the training as well as information about attendance of teachers and administrative staff who were not eligible to participate in the study (such as grade 1 teachers). Overall, approximately 80 percent of year 1 teacher participants and 76 percent of year 2 participants in intervention group schools attended the training. Note that teachers that were hired in intervention schools at the beginning of year 2 did not have the opportunity to receive training prior to implementing the curriculum. Differences in attendance rates across school grades were not statistically significant. Principals from 15 of the 25 intervention group schools attended, and an average of 1.76 ineligible teachers or administrative staff per school also attended the training.

²⁵ To ease the burden of participation – the lessons about applying the character trait at home/in the community were made optional to intervention teachers.

Table 3.1. Lessons in Character training attendance

| Attendee | Number | Percent of total |
|---|-------------------|-------------------|
| Principals or associate principals | 15 | 60.0 |
| Nonparticipating teachers or administrative staff | 1.76 ^a | 68.0 ^b |
| <i>Attendance of year 1 teacher participants</i> | | |
| All teachers | 237 | 80.1 |
| Grade 2 teachers | 83 | 83.0 |
| Grade 3 teachers | 59 | 73.7 |
| Grade 4 teachers | 52 | 86.7 |
| Grade 5 teachers | 43 | 76.8 |
| $\chi^2_{(3)} = 4.5, p\text{-value} = 0.21^c$ | | |
| Number of year 1 participating teachers | 296 | — |
| <i>Attendance of year 2 teacher participants</i> | | |
| All teachers | 184 | 75.7 |
| Grade 2 teachers | 57 | 76.0 |
| Grade 3 teachers | 45 | 71.4 |
| Grade 4 teachers | 46 | 82.1 |
| Grade 5 teachers | 36 | 73.5 |
| $\chi^2_{(3)} = 1.5, p\text{-value} = 0.71^c$ | | |
| Number of year 2 participating teachers | 243 | — |

— is not applicable.

a. Average per school.

b. Percentage of schools in which at least one noneligible teacher or staff member attended training.

c. Chi-squared test for difference in rates across school grade.

Source: Professional development training attendance log data.

Lessons in Character core and supplementary lesson implementation

Table 3.2 provides estimates of the number of LIC core and supplementary (Daily Oral Language with Character and Writing with Character) lessons delivered each implementation year. These estimates were obtained using implementation log data provided by teachers in the intervention group as well as correspondence with teachers. The estimates in the table assume that teachers who did not provide implementation log data or other information about implementation did not implement any lessons.

As described above, the LIC core curriculum consists of 25 lessons, of which participating intervention group teachers were asked to implement 19 during the academic year. Teachers implemented an average of 12.40 core lessons in year 1 and 9.56 lessons in year 2. In year 1, 16 percent of teacher did not implement any lessons, while 40 percent implemented 16 or more. The use of curricular materials declined after year 1. In year 2, 28 percent of teachers did not implement any LIC lessons, and 28 percent implemented 16 or more lessons.

Table 3.2 also indicates that the supplementary curricular materials, which were optional for teachers, were used less frequently than the core materials. In year 1, two-thirds of teachers used at least some of the Daily Oral Language with Character or Writing with Character materials in their classrooms. In year 2, about half of teachers reported using these materials.

Table 3.2. Lessons implemented by grade 2-5 teachers in intervention group schools

| Lessons implemented | Year 1 | | | Year 2 | | |
|---|--------|------------------|-----------------------|--------|------------------|-----------------------|
| | Number | Percent of total | Cumulative percentage | Number | Percent of total | Cumulative percentage |
| <i>Lessons in Character core curriculum</i> | | | | | | |
| Mean | 12.40 | | | 9.56 | | |
| (standard deviation) | (8.10) | | | (8.63) | | |
| 0 | 47 | 15.9 | 15.9 | 69 | 28.4 | 28.4 |
| 1-5 | 30 | 10.1 | 26.0 | 31 | 12.8 | 41.1 |
| 6-10 | 40 | 13.5 | 39.5 | 36 | 14.8 | 56.0 |
| 11-15 | 60 | 20.3 | 59.8 | 38 | 15.6 | 71.6 |
| 16-18 | 29 | 9.8 | 69.6 | 14 | 5.8 | 77.4 |
| 19-25 | 90 | 30.4 | 100.0 | 55 | 22.6 | 100.0 |
| <i>Daily Oral Language with Character</i> | | | | | | |
| 0 | 97 | 32.8 | 32.8 | 118 | 48.6 | 48.6 |
| Some | 199 | 67.2 | 100.0 | — | — | — |
| 1-25 | — | — | — | 86 | 35.4 | 84.0 |
| 26-90 | — | — | — | 33 | 13.6 | 97.5 |
| 91-180 | — | — | — | 6 | 2.5 | 100.0 |
| <i>Writing with Character^a</i> | | | | | | |
| 0 | 70 | 35.7 | 35.7 | 83 | 49.4 | 48.6 |
| Some | 126 | 64.3 | 100.0 | — | — | — |
| 1-10 | — | — | — | 53 | 31.5 | 80.9 |
| 11-25 | — | — | — | 28 | 16.7 | 97.6 |
| 26-36 | — | — | — | 4 | 2.4 | 100.0 |
| <i>Implemented any component of LIC</i> | | | | | | |
| No | 42 | 14.2 | 14.2 | 57 | 23.5 | 23.5 |
| Yes | 254 | 85.8 | 100.0 | 186 | 76.5 | 100.0 |

— is not applicable because the number of lessons implemented could not be estimated for cohort 1 teachers in year 1.

Note: Estimates are based on the implementation logs or correspondence received by participating teachers. If no implementation logs or other information was received, it is assumed that no lessons were implemented.

a. Grades 3-5 only.

Source: Teacher implementation log and teacher correspondence data.

Table 3.3 presents estimates of the number of LIC core and supplementary lessons delivered in each implementation year by teachers in grades 4-5, teachers who served students in the primary and exploratory analytic samples. Overall, teachers in grades 4-5 reported implementing similar numbers of lessons as teachers in grades 2-5. Grade 4-5 teachers implemented an average of 12.15 core lessons in year 1 and 8.90 lessons in year 2. Fifteen percent of these teachers did not implement any core lessons in year 1, while 40 percent implemented 16 or more. As was the case for the grade 2-5 sample, the use of curricular materials declined after year 1: 29 percent of teachers did not implement any LIC lessons, and 25 percent implemented 16 or more lessons. Between 62 and 67 percent of grade 4-5 teachers used at least some of the Daily Oral Language with Character or Writing with Character materials in their classrooms in year 1, and about half of teachers reported using these materials in year 2.

Table 3.3. Lessons implemented by grade 4-5 teachers in intervention group schools

| Lessons implemented | Year 1 | | | Year 2 | | |
|---|--------|------------------|-----------------------|--------|------------------|-----------------------|
| | Number | Percent of total | Cumulative percentage | Number | Percent of total | Cumulative percentage |
| <i>Lessons in Character core curriculum</i> | | | | | | |
| Mean | 12.15 | | | 8.90 | | |
| (standard deviation) | (8.22) | | | (8.25) | | |
| 0 | 17 | 14.7 | 14.7 | 30 | 28.6 | 28.6 |
| 1-5 | 17 | 14.7 | 29.3 | 18 | 17.1 | 45.7 |
| 6-10 | 13 | 11.2 | 40.5 | 11 | 10.5 | 56.2 |
| 11-15 | 23 | 19.8 | 60.3 | 20 | 19.0 | 75.2 |
| 16-18 | 12 | 10.3 | 70.7 | 5 | 4.8 | 80.0 |
| 19-25 | 34 | 29.3 | 100.0 | 21 | 20.0 | 100.0 |
| <i>Daily Oral Language with Character</i> | | | | | | |
| 0 | 38 | 32.8 | 32.8 | 51 | 48.6 | 48.6 |
| Some | 78 | 67.2 | 100.0 | — | — | — |
| 1-25 | — | — | — | 37 | 35.2 | 83.8 |
| 26-90 | — | — | — | 14 | 13.3 | 97.1 |
| 91-180 | — | — | — | 3 | 2.9 | 100.0 |
| <i>Writing with Character</i> | | | | | | |
| 0 | 44 | 37.9 | 37.9 | 53 | 50.5 | 50.5 |
| Some | 72 | 62.1 | 100.0 | — | — | — |
| 1-10 | — | — | — | 33 | 31.4 | 81.9 |
| 11-25 | — | — | — | 16 | 15.2 | 97.1 |
| 26-36 | — | — | — | 3 | 2.9 | 100.0 |
| <i>Implemented any component of LIC</i> | | | | | | |
| No | 16 | 13.8 | 13.8 | 24 | 22.9 | 22.9 |
| Yes | 100 | 86.2 | 100.0 | 81 | 77.1 | 100.0 |

— is not applicable because the number of lessons implemented could not be estimated for cohort 1 teachers in year 1.

Note: Estimates are based on the implementation logs or correspondence received by participating teachers. If no implementation logs or other information was received, it is assumed that no lessons were implemented.

Source: Teacher implementation log and teacher correspondence data.

Among teachers who did implement core or supplementary lessons, most reported that they used the materials as instructed in the implementation guide. Some 53–75 percent of teachers reported using the materials “mostly” or “exactly” as instructed in the implementation guide (table 3.4).

Among grade 4-5 teachers who implemented the lessons, between 44 and 75 percent reported that they used the materials “mostly” or “exactly” as instructed in the implementation guide, depending on the year and curriculum type (table 3.5).

Table 3.4. Teacher-reported implementation fidelity in intervention group schools (grade 2-5 teachers)

| Curriculum type | Year 1 | | | Year 2 | | |
|---|--------|---------------------------------|-----------------------|--------|---------------------------------|-----------------------|
| | Number | Percent of total implementation | Cumulative percentage | Number | Percent of total implementation | Cumulative percentage |
| <i>Lessons in Character core curriculum</i> | | | | | | |
| No lessons implemented | 47 | — | — | 69 | — | — |
| Not at all as written | 8 | 3.7 | 3.7 | 9 | 5.4 | 5.4 |
| Sometimes as written | 54 | 24.8 | 28.4 | 52 | 31.5 | 37.0 |
| Mostly or exactly as written | 156 | 71.5 | 100.0 | 104 | 63.0 | 100.0 |
| No response | 31 | — | — | 9 | — | — |
| <i>Daily Oral Language with Character</i> | | | | | | |
| No lessons implemented | 97 | — | — | 118 | — | — |
| Not at all as written | 12 | 6.3 | 6.3 | 11 | 9.3 | 4.1 |
| Sometimes as written | 36 | 19.0 | 25.4 | 41 | 34.7 | 28.8 |
| Mostly or exactly as written | 141 | 74.6 | 100.0 | 66 | 55.9 | 100.0 |
| No response | 10 | — | — | 7 | — | — |
| <i>Writing with Character^a</i> | | | | | | |
| No lessons implemented | 70 | — | — | 83 | — | — |
| Not at all as written | 11 | 9.4 | 9.4 | 7 | 8.6 | 8.6 |
| Sometimes as written | 37 | 31.6 | 41.0 | 31 | 38.3 | 46.9 |
| Mostly or exactly as written | 69 | 58.9 | 100.0 | 43 | 53.0 | 100.0 |
| No response | 9 | — | — | 4 | — | — |

— is not applicable

Note: Data are median self-reported implementation across implementation logs submitted during the academic year.

a. Grades 3–5 only.

Source: Teacher implementation log data.

Table 3.5. Teacher-reported implementation fidelity in intervention group schools (grade 4-5 teachers)

| Curriculum type | Year 1 | | | Year 2 | | |
|---|--------|---------------------------------|-----------------------|--------|---------------------------------|-----------------------|
| | Number | Percent of total implementation | Cumulative percentage | Number | Percent of total implementation | Cumulative percentage |
| <i>Lessons in Character core curriculum</i> | | | | | | |
| No lessons implemented | 17 | — | — | 30 | — | — |
| Not at all as written | 4 | 4.6 | 4.6 | 6 | 8.7 | 8.7 |
| Sometimes as written | 18 | 20.7 | 25.3 | 22 | 31.9 | 40.6 |
| Mostly or exactly as written | 65 | 74.7 | 100.0 | 41 | 59.4 | 100.0 |
| No response | 12 | — | — | 6 | — | — |
| <i>Daily Oral Language with Character</i> | | | | | | |
| No lessons implemented | 38 | — | — | 51 | — | — |
| Not at all as written | 5 | 6.7 | 6.7 | 7 | 14.3 | 14.3 |
| Sometimes as written | 16 | 21.3 | 28.0 | 19 | 38.8 | 53.1 |
| Mostly or exactly as written | 54 | 72.0 | 100.0 | 23 | 46.9 | 100.0 |
| No response | 3 | — | — | 5 | — | — |
| <i>Writing with Character</i> | | | | | | |
| No lessons implemented | 44 | — | — | 53 | — | — |
| Not at all as written | 6 | 8.7 | 9.4 | 5 | 10.4 | 10.4 |
| Sometimes as written | 28 | 40.6 | 41.0 | 22 | 45.8 | 56.2 |
| Mostly or exactly as written | 35 | 50.7 | 100.0 | 21 | 43.7 | 100.0 |
| No response | 3 | — | — | 4 | — | — |

— is not applicable

Note: Data are median self-reported implementation across implementation logs submitted during the academic year.

Source: Teacher implementation log data.

Evidence of character development activities and artifacts

As described in chapter 2, a site observation checklist was used to assess evidence of social and character development activities and artifacts in public areas and grade 4 and 5 classrooms. Character development banners and posters were the most commonly observed artifact observed in public areas of study schools; they were observed in 52–60 percent of intervention group schools and 43–65 percent of control group schools (table 3.6). Such posters were also commonly observed in study classrooms (49–70 percent), as were student behavior codes (48–66 percent). These results suggest that there are no consistent differences across intervention and control group schools in observed character development-related artifacts or observer ratings.

Informal character development programs and activities

Teachers reported on informal social development activities in the year prior to LIC implementation (teachers only) and during the spring of both implementation years. Most teachers reported activities aimed at promoting students' social development (tables 3.7 and 3.8). There were no statistically significant differences across intervention and control group schools in the percentages of teachers reporting the use of social and character development activities.

Summary of implementation findings

LIC is a literature-based, supplementary language arts program that focuses explicitly on trustworthiness, respect, responsibility, fairness, caring, and citizenship. The curricular material comprises two components: the LIC core curriculum and support materials—Daily Oral Language with Character and Writing with Character—that are designed to support language arts learning standards and reinforce the traits of good character.

As designed, the LIC core curriculum consists of 25 lessons (24 core principal lessons and one STAR lesson), of which intervention group teachers were asked to implement 19 during the academic year. Teachers in intervention schools implemented fewer lessons than they were asked to implement – an average of 12.40 LIC lessons in Year 1 and 9.56 LIC lessons in Year 2. The supplementary curricular materials were used less frequently than the LIC core materials: two-thirds of teachers used at least some of the Daily Oral Language with Character or Writing with Character materials in their classrooms in year 1 and about half of teachers reported using these materials in year 2. In sum, participating teachers implemented fewer core LIC lessons and used the Daily Oral Language with Character and Writing with Character materials less frequently in their classes than recommended by the program developer.

Table 3.6. Observed artifacts indicative of social and character education programs in intervention and control group schools (proctor observation checklist)

| Artifact | Year 1 | | | | Year 2 | | | |
|---|--------------|-------------|------------|----------------------|--------------|-------------|------------|----------------------|
| | Intervention | Control | Difference | p-value ^a | Intervention | Control | Difference | p-value ^a |
| <i>Outside school/front office</i> | | | | | | | | |
| Mean number of artifacts (standard deviation) | 1.60 (1.51) | 1.63 (1.55) | -0.03 | 0.95 | 1.94 (1.70) | 2.11 (1.55) | -0.17 | 0.71 |
| Recognition of character development activities | 30.0% | 23.3% | 6.7 | 0.55 | 31.2% | 34.0% | -2.8 | 0.81 |
| Students recognized for character development | 22.0% | 26.7% | -4.7 | 0.67 | 33.3% | 34.0% | -0.7 | 0.96 |
| Display area focused on character development | 30.0% | 39.7% | -9.7 | 0.45 | 41.7% | 44.0% | -2.3 | 0.85 |
| Character development banners/posters | 52.0% | 42.7% | 9.3 | 0.48 | 60.4% | 65.3% | -4.9 | 0.71 |
| Student character development work displayed | 26.0% | 30.3% | -4.3 | 0.71 | 27.1% | 34.0% | -6.9 | 0.56 |
| Number of schools (observations) | 39 | 42 | | | 36 | 42 | | |
| <i>Classrooms (grades 4-5)</i> | | | | | | | | |
| Mean number of artifacts (standard deviation) | 0.83 (0.71) | 0.66 (0.57) | 0.16 | 0.38 | 1.05 (0.71) | 1.06 (0.69) | -0.02 | 0.92 |
| Recognition of character development activities | 22.0% | 18.8% | 3.2 | 0.73 | 32.7% | 33.9% | -1.2 | 0.91 |
| Students recognized for character development | 20.4% | 16.1% | 4.3 | 0.60 | 30.3% | 38.4% | -8.1 | 0.44 |
| Display area focused on character development | 38.9% | 29.2% | 9.7 | 0.36 | 42.0% | 44.7% | -2.7 | 0.79 |
| Character development banners/posters | 55.1% | 48.7% | 6.4 | 0.56 | 55.3% | 70.0% | -14.7 | 0.12 |
| Student character development work displayed | 17.0% | 27.5% | -10.5 | 0.22 | 28.9% | 33.8% | -4.9 | 0.64 |
| Character education curriculum materials | 22.3% | 13.6% | 8.7 | 0.30 | 34.0% | 34.6% | -0.5 | 0.96 |
| Student behavior code | 55.5% | 48.1% | 7.4 | 0.49 | 62.6% | 65.8% | -3.3 | 0.73 |

| Artifact | Year 1 | | | | Year 2 | | | |
|--|--------------|---------|------------|----------------------|--------------|---------|------------|----------------------|
| | Intervention | Control | Difference | p-value ^a | Intervention | Control | Difference | p-value ^a |
| Positive behavior chart | 36.7% | 31.3% | 5.4 | 0.57 | 58.6% | 44.5% | 14.1 | 0.16 |
| Number of classrooms | 119 | 113 | | | 115 | 125 | | |
| <i>Observer global ratings</i> | | | | | | | | |
| How clearly does school represent a focus on social and character development? | | | | | | | | |
| Not at all | 9.5 | 6.0 | 3.5 | 0.59 ^b | 6.2 | 6.0 | 0.2 | 0.53 ^b |
| A little | 21.4 | 13.3 | 8.1 | | 29.2 | 17.3 | 11.9 | |
| Somewhat | 14.3 | 22.7 | -8.4 | | 27.1 | 32.0 | -4.9 | |
| A lot | 28.6 | 31.3 | -2.7 | | 22.9 | 30.0 | -7.1 | |
| Very much | 26.7 | 26.2 | 0.5 | | 14.6 | 14.7 | -0.1 | |
| Visuals (such as posters) related to good character throughout the school? | | | | 0.39 ^b | | | | 0.26 ^b |
| None | 20.5 | 10.0 | 10.5 | | 16.7 | 9.3 | 7.4 | |
| A few | 25.5 | 27.3 | -1.8 | | 29.2 | 33.3 | -4.1 | |
| Some | 38.6 | 43.3 | -4.7 | | 41.7 | 25.3 | 16.4 | |
| A lot | 15.9 | 19.3 | -3.4 | | 12.5 | 32.0 | -19.5 | |
| Number of observations | 39 | 41 | | | 36 | 42 | | |

Note: Data are school averages of observer ratings.

a. An independent samples *t*-test was used to estimate intervention-control group differences in school averages of observer ratings. No multiple comparison adjustments were applied. No differences were statistically significant at the 0.05 level (two-tail).

b. Statistical test is based on the school average of the item. The item was treated as if it was measured on an interval scale.

Source: Proctor site observation data.

Table 3.7. Social development activities implemented in study classrooms and schools (grades 2-5, past six months)

| Activity | Intervention (percent) | Control (percent) | Difference (percentage points) | z-test^a | p-value |
|--------------------------|----------------------------------|-----------------------------|---|---------------------------|----------------|
| | | | | | |
| <i>Baseline (Year 0)</i> | | | | | |
| Character education | 81.8 | 78.7 | 3.1 | 0.19 | 0.85 |
| Violence prevention | 64.3 | 58.6 | 5.7 | 0.31 | 0.76 |
| Tolerance and diversity | 72.6 | 62.3 | 10.3 | 0.68 | 0.50 |
| Civics/citizenship | 72.0 | 64.8 | 7.2 | 0.51 | 0.61 |
| Self-discipline | 81.2 | 75.7 | 5.5 | 0.43 | 0.67 |
| Resisting peer pressure | 65.8 | 55.0 | 10.8 | 0.71 | 0.48 |
| Empathy | 73.7 | 58.4 | 15.3 | 0.88 | 0.38 |
| Conflict resolution | 69.5 | 63.6 | 5.9 | 0.38 | 0.70 |
| Behavior management | 87.1 | 82.6 | 4.5 | 0.39 | 0.70 |
| Number of teachers | 289 | 316 | | | |
| <i>Year 1</i> | | | | | |
| Character education | 96.8 | 86.0 | 10.8 | 0.87 | 0.39 |
| Violence prevention | 65.3 | 65.1 | 0.2 | -0.26 | 0.80 |
| Tolerance and diversity | 70.0 | 71.0 | -1.0 | -0.07 | 0.94 |
| Civics/citizenship | 80.7 | 68.0 | 12.7 | 0.80 | 0.42 |
| Self-discipline | 83.8 | 77.2 | 6.6 | 0.48 | 0.63 |
| Resisting peer pressure | 68.6 | 52.1 | 16.5 | 1.06 | 0.29 |
| Empathy | 71.4 | 62.5 | 8.9 | 0.62 | 0.54 |
| Conflict resolution | 82.1 | 66.0 | 16.1 | 0.79 | 0.43 |
| Behavior management | 85.0 | 82.2 | 2.8 | 0.19 | 0.85 |
| Number of teachers | 252 | 310 | | | |
| <i>Year 2</i> | | | | | |
| Character education | 90.1 | 76.3 | 13.7 | 0.83 | 0.40 |
| Violence prevention | 63.5 | 57.9 | 5.4 | 0.27 | 0.79 |
| Tolerance and diversity | 71.6 | 70.2 | 1.2 | 0.09 | 0.93 |
| Civics/citizenship | 71.0 | 72.1 | -1.2 | -0.15 | 0.88 |
| Self-discipline | 77.4 | 77.2 | 0.2 | -0.03 | 0.97 |
| Resisting peer pressure | 67.2 | 57.5 | 9.6 | 0.52 | 0.61 |
| Empathy | 71.0 | 67.6 | 3.3 | 0.17 | 0.87 |
| Conflict resolution | 72.2 | 66.7 | 5.4 | 0.37 | 0.71 |
| Behavior management | 77.8 | 83.0 | -5.2 | -0.45 | 0.65 |
| Number of teachers | 208 | 295 | | | |

Note: Data are based on teacher reports.

a. A two-tailed test that adjusted for intraclass correlations was used to test differences in reported social development activities across intervention and control schools. In no case was a statistically significant difference found. No multiple comparison adjustments were applied.

Source: Baseline, year 1, and year 2 teacher survey data.

Table 3.8. Social development activities implemented in study classrooms and schools (grades 4-5, past six months)

| Activity | Intervention (percent) | Control (percent) | Difference (percentage points) | <i>z</i> -test ^a | <i>p</i> -value |
|--------------------------|----------------------------------|-----------------------------|---|-----------------------------|-----------------|
| | | | | | |
| <i>Baseline (Year 0)</i> | | | | | |
| Character education | 77.3 | 74.6 | 2.7 | 0.41 | 0.68 |
| Violence prevention | 63.8 | 57.9 | 5.9 | 0.43 | 0.66 |
| Tolerance and diversity | 68.8 | 60.8 | 8.0 | 0.60 | 0.55 |
| Civics/citizenship | 67.5 | 57.6 | 9.9 | 0.69 | 0.49 |
| Self-discipline | 77.2 | 74.9 | 2.3 | 0.19 | 0.85 |
| Resisting peer pressure | 72.0 | 61.1 | 10.9 | 0.79 | 0.43 |
| Empathy | 70.2 | 57.3 | 12.9 | 0.91 | 0.36 |
| Conflict resolution | 68.9 | 63.9 | 5.0 | 0.35 | 0.73 |
| Behavior management | 83.9 | 76.3 | 7.6 | 0.65 | 0.72 |
| Number of teachers | 115 | 121 | | | |
| <i>Year 1</i> | | | | | |
| Character education | 97.8 | 80.4 | 17.4 | 1.64 | 0.10 |
| Violence prevention | 71.5 | 62.4 | 9.1 | 0.52 | 0.60 |
| Tolerance and diversity | 73.3 | 70.3 | 3.0 | 0.22 | 0.82 |
| Civics/citizenship | 81.8 | 66.0 | 15.8 | 1.20 | 0.23 |
| Self-discipline | 87.3 | 77.8 | 9.5 | 0.61 | 0.54 |
| Resisting peer pressure | 72.6 | 57.8 | 14.8 | 1.02 | 0.31 |
| Empathy | 70.6 | 56.1 | 14.5 | 0.98 | 0.33 |
| Conflict resolution | 81.5 | 66.5 | 15.0 | 0.70 | 0.48 |
| Behavior management | 88.1 | 80.0 | 8.1 | 0.73 | 0.46 |
| Number of teachers | 116 | 94 | | | |
| <i>Year 2</i> | | | | | |
| Character education | 88.6 | 76.6 | 12.0 | 0.83 | 0.41 |
| Violence prevention | 66.9 | 51.0 | 15.9 | 0.86 | 0.39 |
| Tolerance and diversity | 71.7 | 72.1 | -0.4 | -0.05 | 0.96 |
| Civics/citizenship | 68.8 | 63.1 | 5.7 | 0.13 | 0.90 |
| Self-discipline | 75.0 | 70.3 | 4.7 | 0.18 | 0.86 |
| Resisting peer pressure | 69.6 | 56.1 | 13.5 | 0.75 | 0.45 |
| Empathy | 71.6 | 60.1 | 11.5 | 0.78 | 0.43 |
| Conflict resolution | 72.8 | 58.0 | 14.8 | 1.02 | 0.31 |
| Behavior management | 72.0 | 74.5 | -2.5 | -0.23 | 0.82 |
| Number of teachers | 86 | 121 | | | |

Note: Data are based on teacher reports.

a. A two-tailed test that adjusted for intraclass correlations was used to test differences in reported social development activities across intervention and control schools. In no case was a statistically significant difference found. No multiple comparison adjustments were applied.

Source: Baseline, year 1, and year 2 teacher survey data.

4. Impact results

The primary aims of the Lessons in Character (LIC) program are to improve student problem solving skills; to promote student understanding, endorsement, and behavioral enactment of core values; and to boost student academic skills, particularly those related to English language arts. By improving problem solving skills and promoting enactment of character traits, the program is posited to increase social competence and reduce antisocial behavior, and by providing opportunities for students to practice language, grammar, mechanics (spelling, punctuation), and composition, the program is expected to have direct effects on student academic performance. Moreover, LIC aims to enhance skills and habits necessary for academic achievement, such as accountability and perseverance.

To investigate the impacts of the LIC program after the second year of program implementation, student outcomes from the spring of year 2 were examined. The analytic sample was restricted to grade 4 and 5 student participants who were in study schools in year 1 and year 2. Second-year impacts on teacher-reported classroom outcomes (belongingness and school expectations) were estimated using data collected from teachers of grades 2–5 in year 2.

Summary of impacts of Lessons in Character

Overall, there were no statistically significant LIC impacts on grade 4 and 5 students' measures of academic achievement, social competence, or problem behavior after the second year of program implementation. None of the estimated intervention-control group school differences in adjusted mean scores were statistically significant. Moreover, the confirmatory intermediate impact analyses indicated that there were no statistically significant LIC impacts on the school environment measures, whether measures are based on student reports or teacher reports.

Primary confirmatory impacts of Lessons in Character (student outcomes)

As described in chapters 1 and 2, three primary student outcome domains were delineated to reflect key outcome areas—academic performance, social competence, and problem behaviors—and impact analyses were conducted for two to three outcomes within each domain. Evidence supporting the effectiveness of the intervention on an outcome within a domain was provided if statistically significant, beneficial impacts were detected on any one of the measures within the outcome area after adjusting for multiple hypotheses.

Primary research question 1

1. Do students in grades 4 and 5 who attend schools in the LIC intervention group exhibit higher scores on measures of academic achievement after two academic years of potential LIC exposure than their counterparts who attend schools in the control group?

The intervention did not produce statistically significant gains on the measures of academic achievement after the second year of program implementation (table 4.1). On average, students in intervention group schools scored 0.090 standard deviation units higher on English language arts tests and 0.078 standard deviation units higher on teachers' reports of academic competence

than their counterparts in control group schools. Neither of these differences is statistically significant at the 0.05 level.

Primary research question 2

2. Do students in grades 4 and 5 who attend schools in the LIC intervention group exhibit higher scores on measures of social competence after two academic years of potential LIC exposure than their counterparts who attend schools in the control group?

The intervention did not impact grade 4 and 5 measures of social competence after two academic years of program implementation. Students in intervention group schools scored 0.129 standard deviation units higher on social skills than students in control group schools. They also scored 0.097 standard deviation units lower on self-reports of altruism, and 0.034 standard deviation lower on self-reports of empathy (see table 4.1). None of the estimated impacts on the social competence measures is statistically significant at the 0.05 level.

Primary research question 3

3. Do students in grades 4 and 5 who attend schools in the LIC intervention group exhibit fewer problem behaviors after two academic years of potential LIC exposure than their counterparts who attend schools in the control group?

Estimated impacts on the measures of student problem behaviors were not statistically significant (see table 4.1). Students in intervention group schools scored, on average, 0.153 standard deviations lower on externalizing behavior but 0.044 standard deviations higher on aggression than their counterparts in control group schools. Estimated intervention-control group school differences in self-reported delinquency were almost nonexistent (effect size = 0.000).

Table 4.1. Impact analysis of primary outcome variables, primary analytic sample second-year impacts

| Impact measure | Adjusted means | | | | | |
|-------------------------------------|--------------------------------------|---------------------------------|--------------------------------|--------------------------------------|-------------|---------------------|
| | Intervention (standard deviation) | Control (standard deviation) | Difference (standard error) | p-value (unadjusted) ^a | Effect size | Student sample size |
| Academic achievement | | | | | | |
| English language arts test scores | 0.075 (0.925) | -0.011 (0.949) | 0.085 (0.051) | 0.192 (0.096) | 0.090 | 4,525 |
| Academic competence | 31.810 (9.534) | 31.042 (9.810) | 0.769 (0.774) | 0.642 (0.321) | 0.078 | 4,376 |
| Social competence | | | | | | |
| Social Skills Rating System | 46.964 (12.448) | 45.363 (12.431) | 1.600 (1.470) | 0.834 (0.278) | 0.129 | 4,393 |
| Altruism | 4.163 (0.737) | 4.231 (0.694) | -0.068 (0.043) | 0.342 (0.114) | -0.097 | 4,511 |
| Empathy | 2.225 (0.409) | 2.238 (0.384) | -0.013 (0.024) | 0.999 (0.588) | -0.034 | 4,512 |
| Problem behavior | | | | | | |
| Externalizing behavior ^b | 1.557 (2.518) | 1.949 (2.567) | -0.392 (0.351) | 0.804 (0.268) | -0.153 | 4,379 |
| Aggression ^b | 1.373 (0.542) | 1.350 (0.516) | 0.023 (0.045) | 0.999 (0.609) | 0.044 | 4,507 |
| Delinquent behavior ^b | 1.176 (0.339) | 1.176 (0.319) | 0.000 (0.026) | 0.999 (0.996) | 0.000 | 4,507 |

Note: Data are regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable.

a. A Bonferroni procedure was used to calculate adjusted p-values across the three composite outcome variables. The adjusted p-value was calculated by multiplying the unadjusted p-value by either 2 (academic performance measures) or 3 (social competence and problem behavior measures (truncated at 0.999).

b. Coded such that higher values corresponding to higher levels of problem behaviors.

Source: Baseline and year 2 teacher-reported Social Skills Rating System and district archival data and year 2 student survey data.

Primary intermediate impacts of Lessons in Character (school environment)

As described in chapter 1, the theory of action that links the LIC program to students' academic achievement, social competence, and problem behaviors holds that the curriculum would impact students both directly and indirectly via the school environment. The curriculum and the posited resulting changes in student behavior are expected to positively affect the school environment by promoting a caring environment and a common set of student expectations, thereby making the school environment more conducive for learning. To examine part of this logic model, the impact of the LIC program on two indicators of the school environment was examined. Because the primary and exploratory student analytic samples are restricted to students in grades 4-5, results based on the teacher surveys are shown for the subsample of grade 4-5 teachers as well as the overall sample of teachers in grades 2-5.

Intermediate research question 4

4. Do teachers and students in the LIC intervention group report greater levels of student belongingness after two years of program implementation than their counterparts in the control group?

The intervention did not impact teacher or student reports of student belongingness. Intervention-control group school differences in adjusted mean scores on the measures of student belonging were not statistically significant, and the estimated impacts ranged from -0.044 to 0.130 standard deviations (table 4.2).

Intermediate research question 5

5. Do teachers and students in the LIC intervention group report greater levels of school expectations consistent with character development after two years of program implementation than their counterparts in the control group?

The intervention was not associated with gains in teacher or student reports of school expectations. Adjusted mean scores on the measures of school expectations were not statistically different across intervention and control group schools (table 4.2).

Table 4.2. Impact analysis of intermediate outcome variables, primary intermediate analytic sample second-year impacts

| Impact measure | Adjusted means | | | p-value (unadjusted) ^a | Effect size | Sample size |
|---|--------------------------------------|---------------------------------|--------------------------------|--------------------------------------|-------------|-------------|
| | Intervention (standard deviation) | Control (standard deviation) | Difference (standard error) | | | |
| Grades 2-5 | | | | | | |
| Student belonging (teacher report) | 3.801 (0.510) | 3.727 (0.569) | 0.074 (0.070) | 0.999 (0.289) | 0.130 | 507 |
| School expectations (teacher report) | 4.586 (0.485) | 4.552 (0.479) | 0.034 (0.061) | 0.999 (0.583) | 0.070 | 507 |
| Grades 4-5 | | | | | | |
| Student belonging (teacher report) | 3.827 (0.536) | 3.776 (0.629) | 0.051 (0.101) | 0.999 (0.617) | 0.081 | 209 |
| Student belonging (student report) | 3.244 (0.736) | 3.275 (0.713) | -0.031 (0.066) | 0.999 (0.632) | -0.044 | 4,514 |
| School expectations (teacher report) | 4.574 (0.505) | 4.566 (0.499) | 0.008 (0.084) | 0.999 (0.923) | 0.016 | 209 |
| School expectations (student report) | 4.002 (0.881) | 3.991 (0.842) | 0.011 (0.063) | 0.999 (0.864) | 0.013 | 4,511 |

Note: Data are regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable.

a. A Bonferroni procedure was used to calculate adjusted *p*-values across the four outcome variables. The adjusted *p*-value was calculated by multiplying the unadjusted *p*-value by 6 (truncated at 0.999).

Source: Baseline and year 2 teacher survey, district archival, and year 2 student survey data.

Sensitivity analyses

All the primary analyses were conducted using models that included the full set of relevant covariates described in chapter 2 and for samples in which missing covariate data were imputed using multiple imputation techniques. To examine the robustness of the impact estimates, models were estimated with different combinations of baseline covariates, different methods of handling missing data, different analytic samples, and based on data that were weighted by the inverse of nonresponse subsampling rate for cohort 1 (see appendixes G and H). Impact estimates from the sensitivity analysis models did not differ substantially from those reported elsewhere in this chapter. With one exception, impact estimates from the sensitivity analyses were not statistically significant, and the confidence intervals of all of the estimates overlap considerably. In the case where statistically significant impacts were estimated (three estimated impacts on altruism, table G4), the sign of the effect was negative.²⁶

²⁶ Note that multiple comparison procedures were not used to adjust the significance levels of estimated impacts in the sensitivity analyses.

5. Exploratory results

To guard against false discoveries and to maintain adequate statistical power, the primary confirmatory research questions focused on a limited number of assessed outcome variables and subsamples. As noted in previous chapters, inferences in this report about the effectiveness of the intervention were made only with regards to the subsample and set of measures that were defined as confirmatory primary outcomes during the design stage of the study. To better understand whether and, if so, how program impacts on outcomes change as implementation unfolds, exploratory analyses were conducted to examine potential impacts on the same outcome measures after one year of program implementation. This chapter reports the results of the exploratory research questions on the impact of the Lessons in Character (LIC) program on grade 4 and 5 students' academic achievement, social competence, and problem behaviors after one year of program implementation. First-year outcomes are also presented for the exploratory research questions related to the school environment.

The purpose of the exploratory analyses was to provide suggestive information about the extent to which impacts may or may not have been present after the first implementation year. As described in chapter 3, analyses of LIC implementation data indicated that teachers in intervention schools implemented fewer core and supplementary program lessons in year 2 than in year 1. Thus, although students in grades 4 and 5 in intervention schools over the entire two-year period were likely exposed to more LIC lessons than students in these grades in year 1, students in these grades in year 2 were likely exposed to fewer core and supplementary lessons over the course of the current academic year than were their counterparts in year 1. If the intensity of program implementation during a particular academic year plays a stronger role in impacting the school environment and student outcomes than cumulative exposure to LIC over a two-year period does, impacts might be expected to be more pronounced in year 1 than in year 2.

The results of the exploratory analyses were not used to make inferences about the overall effectiveness of the intervention, but to provide descriptive information that can inform the confirmatory results and be rigorously tested in future studies. Because these analyses were exploratory, no corrections for multiple hypothesis testing were applied. The data-analytic strategies used to conduct the exploratory analyses were analogous to those used to investigate the primary research questions (1–5). To investigate the impacts of student exposure to the LIC program during the first year of program implementation, outcomes for the spring of year 1 were examined, and the analytic sample was restricted to grade 4 and 5 student participants in study schools in the spring of year 1. First-year impacts on classroom outcomes were estimated using data collected from teachers of grades 2–5, teachers of grades 4 and 5, and students in grade 4 and 5 study classrooms in year 1.

Summary of exploratory impacts of Lessons in Character

Similar to the second-year results, the analyses found no statistically significant beneficial LIC impacts on grade 4 or 5 measures of academic achievement, social competence, or problem behaviors after the first year of program implementation. Estimated first-year impacts ranged

from -0.107 to 0.116^{27} standard deviation units across all the measures. A statistically significant LIC impact on one of the eight measures was detected: students in intervention group schools exhibited lower levels of self-reported altruism than their counterparts in control group schools after the first year of program implementation.

The exploratory intermediate impact analyses indicated that there were not statistically significant LIC impacts on the school environment measures after the first year of implementation.

Exploratory impacts of Lessons in Character (student outcomes)

First-year LIC program impacts on the measures of student academic achievement, social competence, and problem behaviors were estimated.

Exploratory research question 6

6. Do students in grades 4 and 5 who attend schools in the intervention group exhibit higher scores on measures of academic achievement and social competence, and fewer problem behaviors after one academic year of potential LIC exposure than their counterparts who attend schools in the control group?

The intervention did not produce statistically significant gains in student academic achievement or social competence after the first year of program implementation, nor was the intervention associated with declines in student problem behaviors (table 5.1). Of the eight analyzed measures, one was impacted by the intervention in a statistically significant manner, albeit in a manner that favored the control group: students in intervention group schools reported significantly lower levels of altruism than students in control group schools ($p < .05$, effect size = -0.107).

²⁷ Although the estimated effect size of externalizing behavior is -0.116 , because externalizing behavior is coded such that higher values correspond to higher levels of externalizing behavior, an effect size of 0.116 is used in describing the range of impact estimates across the measured subdomains.

Table 5.1. Impact analysis of exploratory outcome variables, exploratory analytic sample first-year impacts

| Impact measure | Adjusted means | | Difference (standard error) | p-value | Effect size | Student sample size |
|-------------------------------------|--|---|--|----------------|------------------------|--------------------------------|
| | Intervention (standard deviation) | Control (standard deviation) | | | | |
| Academic achievement | | | | | | |
| English language arts test scores | 0.010 (0.943) | 0.028 (0.930) | -0.018 (0.043) | 0.674 | -0.019 | 5,355 |
| Academic competence | 31.911 (9.715) | 30.717 (10.046) | 0.195 (0.837) | 0.817 | 0.019 | 5,120 |
| Social competence | | | | | | |
| Social Skills Rating System | 45.438 (13.101) | 44.737 (13.205) | 0.702 (3.070) | 0.820 | 0.053 | 5,149 |
| Altruism | 4.171 (0.751) | 4.245 (0.683) | -0.073** (0.031) | 0.016 | -0.107 | 5,410 |
| Empathy | 2.229 (0.403) | 2.262 (0.375) | -0.033 (0.020) | 0.101 | -0.089 | 5,410 |
| Problem behaviors | | | | | | |
| Externalizing behavior ^a | 1.819 (2.840) | 2.140 (2.767) | -0.321 (0.846) | 0.706 | -0.116 | 5,120 |
| Aggression ^a | 1.392 (0.552) | 1.369 (0.531) | 0.023 (0.033) | 0.480 | 0.043 | 5,390 |
| Delinquent behavior ^a | 1.186 (0.333) | 1.173 (0.315) | 0.013 (0.019) | 0.483 | 0.042 | 5,391 |

** Significantly different from zero at the .05 level (unadjusted), two-tailed test.

Note: Data are regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable.

a. Outcome variables were coded such that higher values corresponding to higher levels of problem behaviors.

Source: Baseline and year 1 teacher-reported Social Skills Rating System, student survey, and district archival data.

Exploratory research question 7

7. Do students and teachers in the LIC intervention group report greater levels of student belongingness and higher levels of school expectations after one year of program implementation than their counterparts in the control group?

The intervention did not impact teacher or student reports of student belongingness or school expectations after the first year of implementation. Intervention-control group school differences in adjusted mean scores on the measures of student belonging and school expectations were not statistically significant (table 5.2). Although the difference was not significant at the 0.05 level, intervention group schools exhibited higher scores on grades 2-5 on teacher-reported school expectations (effect size = 0.156; $p < 0.10$) than control group schools.

Table 5.2. Impact analysis of exploratory outcome variables, exploratory analytic sample first-year impacts (grades 2-5)

| Impact measure | Adjusted means | | Difference (standard error) | <i>p</i> -value | Effect size | Student sample size |
|--------------------------------------|-----------------------------------|------------------------------|-----------------------------|-----------------|-------------|---------------------|
| | Intervention (standard deviation) | Control (standard deviation) | | | | |
| Grades 2-5 | | | | | | |
| Student belonging (teacher report) | 3.768 (0.547) | 3.707 (0.577) | 0.061 (0.064) | 0.343 | 0.106 | 564 |
| School expectations (teacher report) | 4.625 (0.456) | 4.549 (0.489) | 0.076 (0.043) | 0.079 | 0.156 | 564 |
| Grades 4-5 | | | | | | |
| Student belonging (teacher report) | 3.738 (0.629) | 3.670 (0.606) | 0.068 (0.064) | 0.506 | 0.112 | 210 |
| Student belonging (student report) | 3.233 (0.724) | 3.223 (0.719) | 0.010 (0.056) | 0.860 | 0.014 | 5,415 |
| School expectations (teacher report) | 4.601 (0.468) | 4.487 (0.519) | 0.114 (0.076) | 0.134 | 0.221 | 210 |
| School expectations (student report) | 3.986 (0.881) | 3.992 (0.863) | -0.006 (0.046) | 0.891 | -0.007 | 5,413 |

Note: Data are regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control group standard deviation of the outcome variable.

Source: Baseline and year 1 teacher survey-, district archival, and year 1 student survey data.

Sensitivity analyses

As was the case for the primary impact estimates presented in chapter 4, models were estimated with different combinations of baseline covariates, different methods of handling missing data, different analytic samples, and with weights applied to examine the stability of the exploratory impact estimates (see appendixes I and J). For all the outcomes considered except for the altruism subdomain and teachers' reports of school expectations, impact estimates from the sensitivity analyses were not statistically significant, and the confidence intervals of all of the estimates overlap considerably. The intervention impacts on student reports of altruism were negative and statistically significant across each of the models and analytic samples, with impact

effect sizes ranging from -0.090 to -0.116 . Estimated intervention impacts on teacher reports of school expectations were positive and similar in magnitude across models and samples. Effect sizes ranged from 0.115 to 0.181 . The estimated impact was statistically significant at the 0.05 level when the missing-indicator method was used to handle missing covariate data, but was not statistically significant for the other models and analytic samples.

6. Summary of study findings and study limitations

Overview of findings

On average, teachers in intervention group schools implemented 12.40 core lessons in year 1 and 9.56 lessons in year 2. In year 1, 16 percent reported that they did not deliver any lessons, while 40 percent delivered 16 or more. Teacher-reported use of the curricular materials declined after year 1. In year 2, 28 percent of teachers did not implement any LIC lessons, and 28 percent implemented 16 or more lessons. The supplementary curricular materials were used less frequently than the core materials. In year 1, two-thirds of teachers used at least some of the Daily Oral Language with Character or Writing with Character materials in their classrooms. In year 2, about half of teachers reported using these materials.

The primary confirmatory impact analyses indicate that grade 4 and 5 students who attended schools in the LIC intervention group did not exhibit higher scores on measures of academic achievement or on measures of social competence after two academic years of potential LIC exposure than grade 4 and 5 students who attended schools in the control group. Nor did intervention group students score lower on measures of problem behaviors. Moreover, the intermediate impact analyses indicated that there were no statistically significant LIC impacts on the school environment measures of school expectations and student feelings of belonging. In addition, although participating teachers in intervention group schools reported implementing fewer core and supplementary lessons during the second year than during the first year, exploratory analyses suggested that there were no statistically significant LIC impacts on grade 4-5 student outcomes or on measured school environment outcomes after the first year of program implementation.

Failure to detect impacts on student or school environment outcomes could have been due to a failure of the intervention design, weak implementation of the program, or methodological limitations of the study. With regard to the first possibility, it could be that, as designed, LIC program activities simply do not work to bring about changes in grade 4 and 5 students' academic achievement, social competence, or problem behaviors. The combination of lessons and activities teaching about core values and decision-making may not work to bring about changes in attitudes, behavior, or academic performance—as specified in the program model.

A second alternative is that the LIC program was not implemented with adequate intensity to bring about changes in student and classroom outcomes. As described in chapter 3, 30 percent of participating teachers in intervention schools in year 1 implemented 19 or more LIC core curriculum lessons, the minimum number of lessons recommended by the developer.

Approximately 23 percent of intervention school teachers implemented 19 or more lessons in year 2. Moreover, observer reports of evidence of social and character development implementation and teacher reports of social and character development activities did not differ across intervention and control schools. Relatively high numbers of teachers in control schools implemented social and character development activities—86 percent reported that they implemented character education activities in year 1 and 76 percent reported that they did so in year 2. The high rates of character education implementation in control schools combined with the apparent low levels of implementation of the LIC core curriculum in intervention schools

suggests that the level of implementation in intervention schools may not have been adequate to generate intervention-control group differences in outcomes.

In sum, the Lessons in Character program—already designed to be a low dosage, easy-to-implement character education program—was implemented at a lower dosage than recommended by the developer. Moreover, standard practice in control schools involved relatively high levels of implementation of social and character development activities. These two factors may have been responsible for the non-significant program impacts on school expectations, student feelings of belonging, social competence, problem behaviors, and academic achievement. This finding that control sites reported implementing activities intended to promote social and character development is consistent with the findings of the Social and Character Development Research Consortium (2010). Finding no evidence that the seven social and character development programs improved student outcomes, the consortium speculated that the intensity of social and character development activities in control sites may have reduced the ability of the evaluated programs to generate impacts. Typical character education-aligned activities in schools may be at a high enough level that the implementation of a low-dosage, literature-based character education program makes little additional difference for student and classroom outcomes.

Study limitations

The study has several limitations that compromise the generalizability of the results. First, the findings are based on volunteer teachers and students in schools that were willing to implement the LIC program and participate in a randomized trial. The findings may differ if schools and teachers were required to participate in the intervention. Moreover, as described above, the impacts estimated only generalize to the LIC program implementation levels of the participating teachers. It is unclear whether the level of implementation in participating intervention schools provides an adequate test of the effectiveness of the intervention design.

Another important limitation is that study retention rates differed between teachers in intervention and control sites. Fewer teachers were retained across years in intervention schools than in control schools, which resulted in lower rates of teacher- and student-provided data in intervention schools than in control schools. At the end of year 2, approximately 70 percent of teachers in study schools provided survey data, compared with 90 percent of teachers in control schools. Such group differences in participation rates could bias impact estimates if participation is associated with outcome measures. There was little evidence of selective participation based on the analyses of baseline differences across intervention and control schools. Although no statistically significant baseline differences in teacher or student characteristics were detected, unobserved differences between intervention and control schools could still cause biases in estimated program impacts.

Appendix A. Statistical power estimates

This appendix presents statistical power estimates conducted during the design stage of the study and subsequent estimates calculated based on the observed data.

Design stage statistical power estimates

In order to determine the appropriate sample sizes required for the study, minimum detectable effect sizes were calculated (see Bloom 1995) based on the unit of randomization, the level of clustering, the availability of baseline explanatory variables, and other design characteristics using the procedures described by Donner and Klar (2000), Murray (1998), Raudenbush (1997), and Schochet (2005). Minimum detectable effect size estimates represent the smallest true program impacts in standard deviation units that can be detected with high probability (Bloom 1995). As defined in the design work, the minimum detectable effect size of a particular study is the smallest effect size that has at least an 80 percent probability of being found statistically significant with 95 percent confidence.

Table A1 presents the parameters used to estimate minimum detectable effect sizes during the planning phase of the study and the parameters used subsequently based on the achieved analytic sample. As discussed in chapter 2, 50 schools were planned to be randomly assigned to two conditions, with an estimated three teachers per grade in each school and 25 students per class. A student attrition rate of about 25 percent was assumed for power estimation purposes, leaving 18 students per class at the end of the second implementation year available for analysis of the outcomes assessed. For the purposes of the power analyses, school-and teacher-level intraclass correlations of 0.15 were assumed for student academic performance and 0.07 for the student nonacademic primary outcomes based on intraclass correlation estimates from group randomized evaluations of student achievement (Hedges and Hedberg 2007; Schochet 2008) and school-based prevention trials (Murray and Hannan 1990; Murray and Blitstein 2003; Murray and Short 1996; Scheier et al. 2002). The intraclass correlations for classroom intermediate outcomes were expected to be 0.15. The statistical power analyses also assumed between- and within-school R^2 values of 0.50 (Schochet 2005). Finally, using a Bonferroni adjustment as a conservative approximation of the proposed resampling method, the critical value of the statistical significance test (0.05) was divided by the number of primary measures analyzed within each outcome area (academic performance, social competence, and problem behaviors). For the intermediate outcome measures, the critical value was divided by the number of intermediate school environment outcomes (6).

Table A1. Parameters used to estimate statistical power in planning phase and based on achieved sample

| Parameter | Planning phase | | Final analytic sample | |
|--|-------------------------|-------------------------|------------------------------|-------------------------|
| | Student outcomes | Teacher outcomes | Student outcomes | Teacher outcomes |
| Schools per condition | 25 | 25 | — ^a | — ^a |
| Participating teachers per school | 6 | 12 | 5.31 | 10.50 |
| Participating students per teacher | 18 | — | 18.01 | — |
| <i>Intraclass correlation</i> | | | | |
| Student academic performance measures | 0.15 | — | 0.19 ^b | — |
| Student nonacademic measures | 0.07 | — | 0.07 ^c | — |
| Teacher-reported—school expectations | — | 0.15 | — | 0.05 |
| Teacher-reported—student belonging | — | 0.15 | — | 0.22 |
| <i>R</i> ² (<i>within-classroom</i>) | | | | |
| Student academic performance measures | 0.50 | — | 0.47 ^d | — |
| Student nonacademic measures | 0.50 | — | 0.09 ^c | — |
| <i>R</i> ² (<i>between-classroom/within-school</i>) | | | | |
| Student academic performance measures | 0.50 | — | 0.40 ^d | — |
| Student nonacademic measures | 0.50 | — | 0.15 ^c | — |
| <i>R</i> ² (<i>within-school</i>) | | | | |
| Teacher-reported classroom outcomes | — | 0.20 | — | 0.15 ^e |
| <i>R</i> ² (<i>between-school</i>) | | | | |
| Student academic performance measures | 0.50 | — | 0.91 ^d | — |
| Student nonacademic measures | 0.50 | 0.20 | 0.50 ^c | — |
| Teacher-reported classroom outcomes | — | 0.20 | — | 0.50 ^e |

— is not applicable.

a. Number suppressed to reduce disclosure risk

b. Maximum intraclass correlation for English language arts test scores

c. Average across 8 nonacademic outcome measures

d. Average across two academic outcome measures

e. Average across two teacher-reported measures

Source Baseline and year 2 teacher-reported Social Skills Rating System, teacher survey, district archival, and year 2 student survey data.

As shown in table A2, with 25 schools per condition, six classrooms per school, and a minimum of 18 students in each classroom, the estimated minimum detectable effect sizes were 0.27 standard deviation for student academic performance, 0.20 standard deviation for social competence and problem behaviors, 0.41 standard deviation for teacher-reported school environment measures, and 0.21 standard deviation for student-reported school environment measures.

Table A2. Minimum detectable effect size estimates for student and teacher outcome measures

| Outcome measure | Planned sample | Achieved sample |
|--|----------------|-----------------|
| Primary outcomes | | |
| <i>Students</i> | | |
| Academic outcomes | 0.27 | 0.19 |
| Nonacademic outcomes | 0.20 | 0.22 |
| Intermediate outcomes | | |
| <i>Student-reported classroom outcomes</i> | 0.21 | 0.23 |
| <i>Teacher-reported classroom outcomes</i> | | |
| School expectations | 0.41 | 0.31 |
| Student belonging | 0.41 | 0.41 |

Note: Calculations assumed type I error rates of .05 (two-sided) and a fixed-effects statistical model. See table A1 for other parameters used to estimate minimum detectable effect sizes.

Source: Computations based on values in table A1.

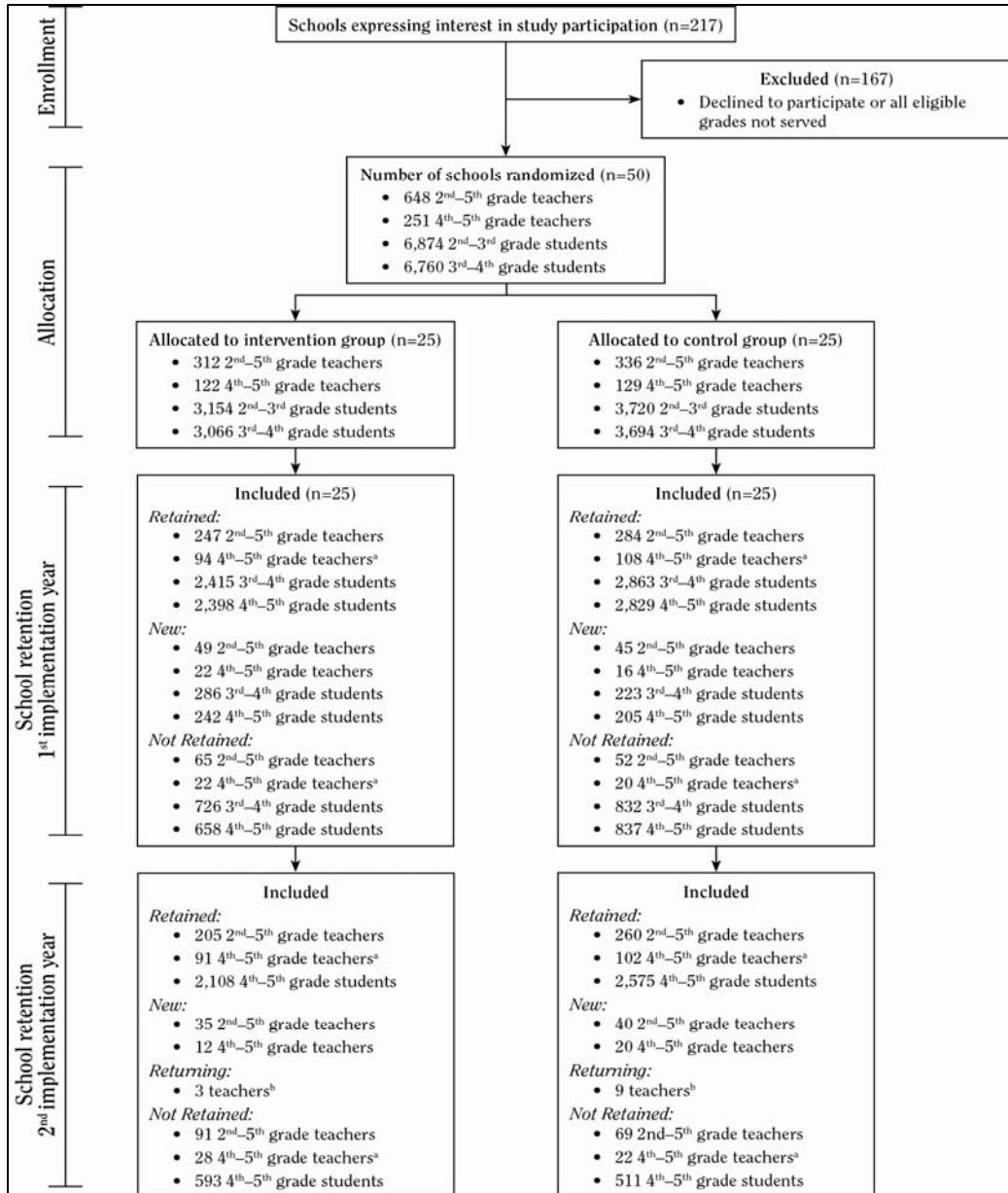
Statistical power estimates based on final analytic sample

As discussed in the report, a small number of schools formally dropped out of the study after the first implementation year. Moreover, fewer teachers participated in each school than anticipated, with an average of 5.31 teacher participants per school. For student academic outcomes, the intraclass correlation was higher than expected (0.19 compared with 0.15), the estimated within-teacher and between-teacher R^2 values were similar to those expected during the design phase, and the between-school R^2 value was substantially higher than anticipated (0.91 compared with 0.50). These differences result in statistical power gains, with a minimum detectable effect size of 0.19 for academic outcomes (see table A2). The minimum detectable effect sizes estimated for student nonacademic outcomes for the final analytic sample were similar to those estimated during the design phase (0.22 compared with 0.20) even though the within-classroom and between-classroom R^2 values were lower than expected. The achieved sample minimum detectable effect size value for teacher-reported school expectations was smaller than originally estimated due to a higher than anticipated between-school R^2 value and the smaller than expected intraclass correlation value. The achieved sample minimum detectable effect size value for teacher-reported student belonging was similar to that estimated during the design phase of the study, as the enhanced statistical power due to a higher than anticipated between-school R^2 was counterbalanced by a reduction in precision due to a higher than expected intraclass correlation value.

To ascertain the robustness of the results, sensitivity analyses were conducted by estimating models that excluded nonintact randomization strata (that is, strata in which schools dropped out of the study) in the analytic sample. After excluding nonintact randomization strata, the estimated minimum detectable effect size was 0.20 standard deviation units for student academic performance, 0.23–0.24 standard deviation units for other student-reported measures, 0.33 standard deviation units for teacher-reported measures of students' feelings of belonging, and 0.34 standard deviation units for teacher-reported measures of school expectations.

Appendix B. Teacher and student recruitment and retention by grade level and data source

Figure B1. Flow diagram of recruitment and retention of schools and study participants

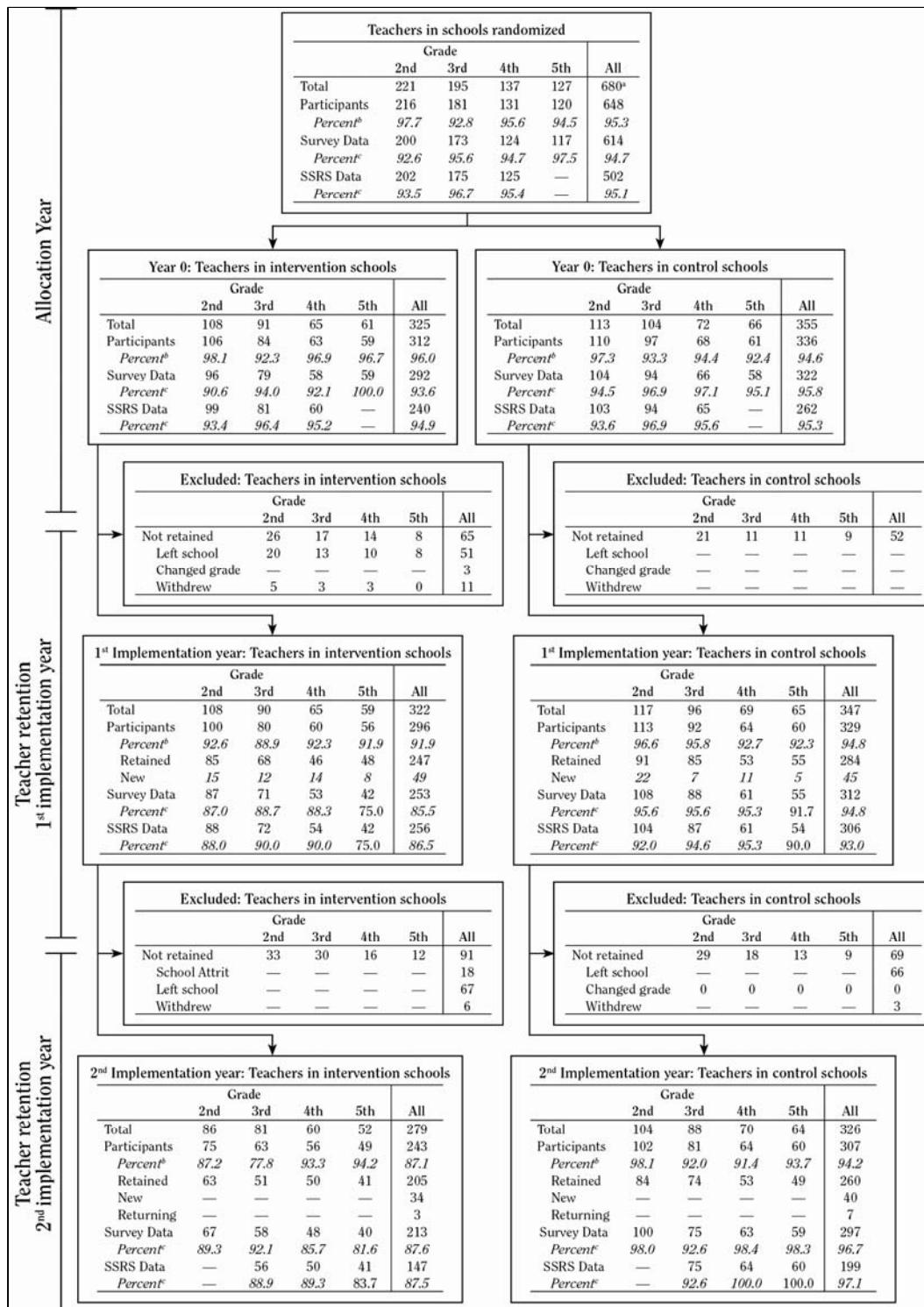


Grade refers to the grade that the teacher taught or that the student was enrolled in during the particular study year (0, 1, 2).

a. The sum of retained and not-retained grade 4 and 5 teachers does not equal the number of included grade 4 and 5 teachers in the previous academic year because some retained teachers taught different grade levels across each year. For example, of the 122 grade 4 and 5 teachers allocated in the intervention group, 22 were not retained in year 1 (122–22=100), 11 taught in grades 2–3 in year 1 (100–11=89), and 5 teachers who taught in grades 2–3 in the allocation year taught in grades 4 or 5 in year 1 (89+5=94).

b. Returning teachers were present in the school in the spring of year 0 or the fall of year 1, not present in the spring of year 1, and present in the spring of year 2.

Figure B2. Flow diagram of teacher recruitment and retention



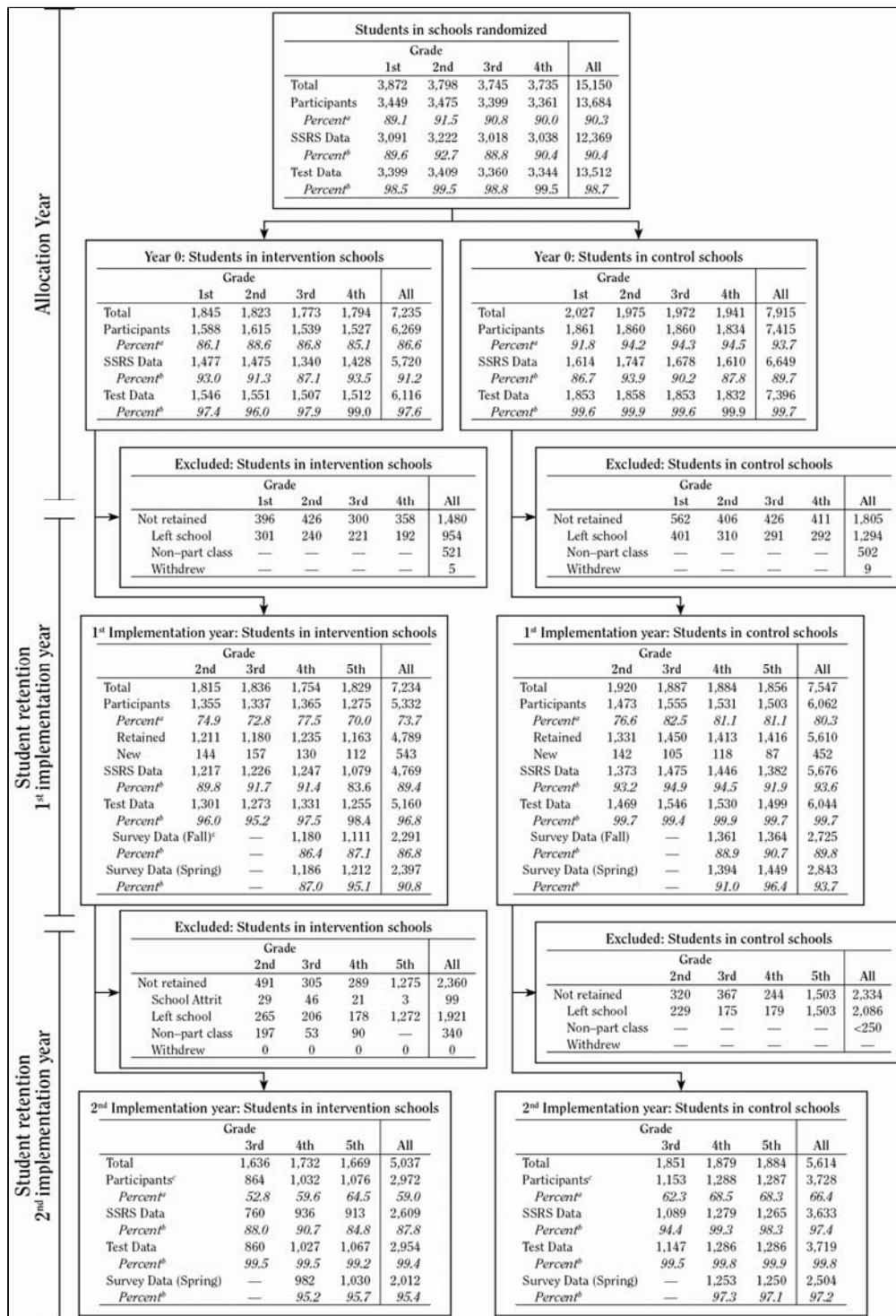
Grade refers to the school grade that the teacher taught during the particular study year (0, 1, 2).

a. Fifteen of the thirty teachers who declined to participate moved out of the school after the spring of year 0.

b. Percentage of eligible teachers in schools participating in study.

c. Percentage of participating teachers returning data.

Figure B3. Flow diagram of student recruitment and retention



Grade refers to the school grade that the student was enrolled in during the particular study year (0, 1, 2).

- Percentage of eligible students in schools participating in study.
- Percentage of participating students returning data.
- All year 2 student participants took part in the study in prior years (that is, no new students were enrolled in year 2 because consent forms were not distributed.)

Appendix C. Survey items used to measure confirmatory and intermediate outcomes

Box C1. Items used to assess primary and exploratory student outcomes

Academic performance

- English language arts test scores
- Academic competence
 - 4 items from the SSRS asking about student academic performance (Gresham and Elliot, 1990)

Social competence

- Social Skills Rating System
 - 30 items from the SSRS assessing cooperation, assertion, and self-control (Gresham and Elliot, 1990)
- Altruism
 - I would help someone learn something
 - Cheer up someone who is feeling sad
 - Get help if I saw a person being picked on
 - Be a friend to someone who is being teased
 - Get help for someone who is hurt
- Empathy
 - When I am mean to someone, I usually feel bad about it later
 - I am happy when the teacher says my friend did a good job
 - I understand how other kids feel
 - Other people's problems really bother me
 - I feel happy when my friend gets a good grade
 - When I see a kid who is upset it really bothers me
 - It's easy for me to tell when my mom or dad has a good day at work
 - It bothers me when my teacher does not feel well
 - I feel sorry for kids who can't find anyone to hang out with
 - Seeing a kid who is crying makes me feel like crying
 - When I see someone who is happy, I feel happy too

Behavior problems

- Externalizing Behavior
 - 6 items from the SSRS assessing externalizing behavior (Gresham and Elliot, 1990)
- Aggression
 - I teased a kid at school
 - I pushed, shoved, or hit a kid from school
 - I called a kid at school a bad name
 - I said that I would hit a kid at school
 - I left out another kid on purpose
 - I made up something about other students to make other kids not like them
- Delinquent Behavior
 - I was sent home from school for bad behavior
 - I was loud or made so much noise at school that I got in trouble
 - I took things from school without paying for them
 - I broke or ruined something on purpose that belonged to the school
 - I took something from school or from the teacher or other kids...
 - I copied other students' homework or copied off the other students' tests
 - I skipped school or didn't go to class without an excuse

Box C2. Items used to assess intermediate and exploratory school environment outcomes**School environment intermediate outcomes^a**

- School expectations
 - The students in this school are expected to get along
 - The students in this school are expected to be kind and caring
 - The students in this school are expected to treat each other fairly
 - The students in this school are expected to obey the rules
 - The students in this school are expected to tell the truth
- Student feelings of belonging
 - The students in this school are nice to each other
 - The students in this school work together on things
 - The students in this school make new students feel welcome
 - The students in this school get along well together
 - The students in this school take good care of school property
 - The students in this school take responsibility for what they do
 - The students in this school respect their teachers
 - The students in this school think it's important to be a good citizen
 - The students in this school treat one another fairly
 - The students in this school tell the truth
 - The students in this school work out problems without fighting

a. School expectations and students feelings of belonging are measured using both teacher and student survey items. Applicable items on the teacher and student surveys are identical.

Appendix D. Internal consistency reliability and intraclass correlations of confirmatory and intermediate outcomes measures

Table D1. Cronbach alpha reliability coefficients and school intraclass correlations of student measures addressing key research questions

| Measure | Number of items | Confirmatory sample | | | | Exploratory sample | | | |
|-----------------------------------|-----------------|---------------------|------------------------------------|----------|------------------------------------|--------------------|------------------------------------|----------|------------------------------------|
| | | Baseline | | Year 2 | | Baseline | | Year 1 | |
| | | α | Intraclass correlation coefficient | α | Intraclass correlation coefficient | α | Intraclass correlation coefficient | α | Intraclass correlation coefficient |
| Academic performance | | | | | | | | | |
| English language arts test scores | a | a | 0.17 | a | 0.18 | a | 0.18 | a | 0.19 |
| Academic competence | 4 | 0.90 | 0.02 | 0.91 | 0.02 | 0.91 | 0.03 | 0.91 | 0.02 |
| Social competence | | | | | | | | | |
| Social Skills Rating System | 30 | 0.95 | 0.04 | 0.96 | 0.08 | 0.96 | 0.04 | 0.97 | 0.05 |
| Altruism | 5 | b | b | 0.83 | 0.05 | 0.79 | 0.02 | 0.93 | 0.03 |
| Empathy | 11 | b | b | 0.78 | 0.06 | 0.72 | 0.02 | 0.77 | 0.04 |
| Behavior problems | | | | | | | | | |
| Externalizing behavior | 6 | 0.90 | 0.03 | 0.90 | 0.05 | 0.91 | 0.02 | 0.91 | 0.02 |
| Aggression | 6 | b | b | 0.84 | 0.06 | 0.82 | 0.04 | 0.84 | 0.04 |
| Delinquent behavior | 7 | b | b | 0.75 | 0.05 | 0.74 | 0.03 | 0.74 | 0.03 |
| School expectations | 5 | b | b | 0.64 | 0.07 | 0.83 | 0.07 | 0.87 | 0.06 |
| Student feelings of belonging | 11 | b | b | 0.88 | 0.13 | 0.85 | 0.09 | 0.87 | 0.11 |

a. Item-level test score data were not collected.

b. Because student surveys were administered to grade 4 and 5 students only, baseline survey data were not available for grade 4 students and not used for grade 5 students in the primary analytic sample.

Source: Baseline, year 1, and year 2 teacher-reported Social Skills Rating System, district archival, student survey data.

Table D2. Cronbach alpha reliability coefficients and intraclass correlations of teacher measures addressing intermediate research questions

| Measure | Number of items | Confirmatory sample | | | | Exploratory sample | | | |
|-------------------------------|-----------------|---------------------|------------------------------------|----------|------------------------------------|--------------------|------------------------------------|----------|------------------------------------|
| | | Baseline | | Year 2 | | Baseline | | Year 2 | |
| | | α | Intraclass correlation coefficient | α | Intraclass correlation coefficient | α | Intraclass correlation coefficient | α | Intraclass correlation coefficient |
| School expectations | 5 | 0.95 | 0.08 | 0.97 | 0.05 | 0.96 | 0.07 | 0.96 | 0.05 |
| Student feelings of belonging | 11 | 0.91 | 0.25 | 0.90 | 0.22 | 0.91 | 0.24 | 0.90 | 0.22 |

Source: Baseline, year 1, and year 2 teacher survey and district archival data.

Appendix E. School-level and teacher-level baseline characteristics for retained sample

Table E1. Teacher baseline demographic characteristics by experimental group

| Characteristic | Baseline percentages | | | Difference (percentage points) | z-test ^a | p-value | | | |
|-----------------------------------|----------------------|------------------------------------|-------------------------------|--------------------------------------|---------------------|---------|--|--|--|
| | Overall (percent) | Intervention group (percent) | Control group (percent) | | | | | | |
| Year 1 Sample | | | | | | | | | |
| <i>Grade taught (year 0)</i> | | | | | | | | | |
| 2 | 34.08 | 33.78 | 34.35 | -0.57 | 0.31 | 0.76 | | | |
| 3 | 27.52 | 27.03 | 27.96 | -0.93 | | | | | |
| 4 | 19.84 | 20.27 | 19.45 | 0.82 | | | | | |
| 5 | 18.56 | 18.92 | 18.24 | 0.68 | | | | | |
| <i>Gender</i> | | | | | | | | | |
| Female | 85.35 | 87.33 | 83.59 | 3.74 | 1.09 | 0.27 | | | |
| Male | 14.65 | 12.67 | 16.41 | -3.74 | | | | | |
| <i>Race/ethnicity^b</i> | | | | | | | | | |
| Asian | 5.89 | 7.01 | 4.84 | 2.17 | 1.13 | 0.26 | | | |
| African American | 7.14 | 8.86 | 5.54 | 3.32 | 1.03 | 0.30 | | | |
| Hispanic | 19.38 | 17.77 | 20.80 | -3.03 | -0.41 | 0.68 | | | |
| White | 85.90 | 83.51 | 88.05 | -4.54 | -1.31 | 0.19 | | | |
| <i>Sample size</i> | | | | | | | | | |
| Schools | 50 | 25 | 25 | | | | | | |
| Teachers ^c | 625 | 296 | 329 | | | | | | |
| Year 2 Sample | | | | | | | | | |
| <i>Grade taught (Year 0)</i> | | | | | | | | | |
| 2 | 32.18 | 30.86 | 33.22 | -2.36 | 0.60 | 0.55 | | | |
| 3 | 26.18 | 25.93 | 26.38 | -0.45 | | | | | |
| 4 | 21.82 | 23.05 | 20.85 | 2.20 | | | | | |
| 5 | 19.82 | 20.16 | 19.54 | 0.62 | | | | | |
| <i>Gender</i> | | | | | | | | | |
| Female | 84.49 | 87.14 | 82.41 | 4.73 | 1.38 | 0.17 | | | |
| Male | 15.51 | 12.86 | 17.59 | -4.73 | | | | | |
| <i>Race/ethnicity^b</i> | | | | | | | | | |
| Asian | 7.40 | 8.88 | 6.18 | 2.70 | 0.86 | 0.39 | | | |
| African American | 6.82 | 8.92 | 5.08 | 3.84 | 1.28 | 0.20 | | | |
| Hispanic | 19.41 | 15.42 | 22.55 | -7.13 | -1.13 | 0.26 | | | |
| White | 86.69 | 85.71 | 87.46 | -1.75 | -0.79 | 0.43 | | | |
| <i>Sample size</i> | | | | | | | | | |
| Teachers ^c | 550 | 243 | 307 | | | | | | |

a. A two-tailed test that adjusted for intraclass correlations was used to test differences in baseline characteristics across intervention and control schools. No multiple comparison adjustments were applied. No differences were statistically significant at the 0.05 level.

b. Racial/ethnic groups are not mutually exclusive.

c. Complete data were available for teacher grade. Approximately 0.6 percent of teacher responses were missing for the gender item, and 1.8 percent of teacher responses were missing for the race/ethnicity item.

Source: Baseline teacher survey and teacher rosters.

Appendix F. Covariates included in impact analysis models

Student-level impact models

As discussed in chapter 2, specific pre-intervention covariates were selected to be included in the impact analysis models based on judgments about the extent to which they would explain variance in the outcome variables and based on study design considerations. All student-level impact analyses included the following covariates: English language arts test scores, grade in school (4 or 5), student gender, English language learner status, race/ethnicity, parent education, dichotomous variables for each randomization stratum, school academic performance (California Academic Performance Index scores), and school-level means of English language test scores in the year prior to program implementation. In addition to these covariates, other pretest covariates were included in the models depending on the outcome measure. For example, academic performance impact analysis models included controls for pre-intervention measures of academic competence. Social competence impact analysis models included controls for pre-intervention scores on the social skills rating scale. Tables G1–G4 display the covariates included in the impact analysis models. More pretest covariates could be included in the exploratory models because of the availability of student-level survey measures assessed prior to program implementation. Pre-intervention student-level survey measures were not available for students who were in grade 4 in the second year of program implementation. Because grade 4 and grade 5 students were pooled in the primary impact analyses, student survey-based pre-intervention covariates were not used for the grade 5 sample. Covariates included in the benchmark models that are the basis of the results in chapters 4 and 5 and the missing-indicator models that are the basis of the sensitivity tests are shown separately.

Teacher-reported intermediate outcome models

All the impact analysis models for intermediate, teacher-reported outcomes included the following covariates: school expectations (teacher report), student belonging (teacher report), grade taught (2–5), teacher gender, years of teaching experience, race/ethnicity, dichotomous indicators for each randomization stratum, school academic performance (California Academic Performance Index scores), and school-level means of year 0 teacher-reports of school expectations and student-belonging. Table G5 presents the covariates included in the teacher-reported intermediate outcome models

Table F1. Covariates included in academic performance impact analysis models

| Variable | Primary analysis (year 2) | | Exploratory analysis (year 1) | |
|--|------------------------------|-------------------------------|----------------------------------|-------------------------------|
| | Benchmark model | Missing indicator model | Benchmark model | Missing indicator model |
| <i>Pretest variables</i> | | | | |
| English language arts test scores (spring year 0) | + | + | + | + |
| Academic competence (spring year 0) | + | + | + | + |
| Social Skills Rating System (spring year 0) | | | | |
| Altruism (fall year 1) | | | | |
| Empathy (fall year 1) | | | | |
| Externalizing behavior (spring year 0) | | | | |
| Aggression (fall year 1) | | | | |
| Delinquent behavior (fall year 1) | | | | |
| School expectations (student, fall year 1) | | | | |
| School belonging (student, fall year 1) | | | | |
| <i>Missing value indicator variables</i> | | | | |
| English language arts test scores | | + | | + |
| Academic competence | | + | | + |
| Social Skills Rating System | | | | |
| Altruism | | | | |
| Empathy | | | | |
| Externalizing behavior | | | | |
| Aggression | | | | |
| Delinquent behavior | | | | |
| School expectations (student) | | | | |
| School belonging (student) | | | | |
| <i>Student demographic variables</i> | | | | |
| Grade (student) | + | + | + | + |
| Gender (student) | + | + | + | + |
| English language learner status | + | + | + | + |
| Race/ethnicity | + | + | + | + |
| Parent education | + | + | + | + |
| <i>School-level covariates</i> | | | | |
| Randomization strata | + | + | + | + |
| California Academic Performance Index scores | + | + | + | + |
| English language arts test scores (grades 2/3, year 0) | + | + | | |
| English language arts test scores (grades 3/4, year 0) | | | + | + |
| Academic competence (grades 2/3, year 0) | + | + | | |
| Academic competence (grades 3/4, year 0) | | | + | + |
| Social skills (grades 2/3, year 0) | | | | |
| Social skills (grades 3/4, year 0) | | | | |
| Altruism (grades 4/5, fall year 1) | | | | |
| Empathy (grades 4/5, fall year 1) | | | | |
| Externalizing behavior (grades 2/3, year 0) | | | | |
| Externalizing behavior (grades 3/4, year 0) | | | | |
| Aggression (grades 4/5, fall year 1) | | | | |
| Delinquent behavior (grades 4/5, fall year 1) | | | | |
| School expectations (student, fall year 1) | | | | |
| School belonging (student, fall year 1) | | | | |

Table F2. Covariates included in social competence impact analysis models

| Variable | Primary analysis (year 2) | | Exploratory analysis (year 1) | |
|--|------------------------------|-------------------------------|----------------------------------|-------------------------------|
| | Benchmark model | Missing indicator model | Benchmark model | Missing indicator model |
| <i>Pretest variables</i> | | | | |
| English language arts test scores (spring year 0) | + | + | + | + |
| Academic competence (spring year 0) | | | | |
| Social Skills Rating System(spring year 0) | + | + | + | + |
| Altruism (fall year 1) | | | + | + |
| Empathy (fall year 1) | | | + | + |
| Externalizing behavior (spring year 0) | | | | |
| Aggression (fall year 1) | | | | |
| Delinquent behavior (fall year 1) | | | | |
| School expectations (student, fall year 1) | | | | |
| School belonging (student, fall year 1) | | | | |
| <i>Missing value indicator variables</i> | | | | |
| English language arts test scores | | + | | + |
| Academic competence | | | | |
| Social Skills Rating System | | + | | + |
| Altruism | | | | + |
| Empathy | | | | + |
| Externalizing behavior | | | | |
| Aggression | | | | |
| Delinquent behavior | | | | |
| School expectations (student) | | | | |
| School belonging (student) | | | | |
| <i>Student demographic variables</i> | | | | |
| Grade (student) | + | + | + | + |
| Gender (student) | + | + | + | + |
| English language learner status | + | + | + | + |
| Race/ethnicity | + | + | + | + |
| Parent education | + | + | + | + |
| <i>School-level covariates</i> | | | | |
| Randomization strata | + | + | + | + |
| California Academic Performance Index scores | + | + | + | + |
| English language arts test scores (grades 2/3, year 0) | + | + | | |
| English language arts test scores (grades 3/4, year 0) | | | + | + |
| Academic competence (grades 2/3, year 0) | | | | |
| Academic competence (grades 3/4, year 0) | | | | |
| Social skills (grades 2/3, year 0) | + | + | | |
| Social skills (grades 3/4, year 0) | | | + | + |
| Altruism (grades 4/5, fall year 1) | + | + | + | + |
| Empathy (grades 4/5, fall year 1) | + | + | + | + |
| Externalizing behavior (grades 2/3, year 0) | | | | |
| Externalizing behavior (grades 3/4, year 0) | | | | |
| Aggression (grades 4/5, fall year 1) | | | | |
| Delinquent behavior (grades 4/5, fall year 1) | | | | |
| School expectations (student, fall year 1) | | | | |
| School belonging (student, fall year 1) | | | | |

Table F3. Covariates included in problem behavior impact analysis models

| Variable | Primary analysis (year 2) | | Exploratory analysis (year 1) | |
|--|--------------------------------------|--|--|--|
| | Benchmark model | Missing indicator model | Benchmark model | Missing indicator model |
| <i>Pretest variables</i> | | | | |
| English language arts test scores (spring year 0) | + | + | + | + |
| Academic competence (spring year 0) | | | | |
| Social Skills Rating System(spring year 0) | | | | |
| Altruism (fall year 1) | | | | |
| Empathy (fall year 1) | | | | |
| Externalizing behavior (spring year 0) | + | + | + | + |
| Aggression (fall year 1) | | | + | + |
| Delinquent behavior (fall year 1) | | | + | + |
| School expectations (student, fall year 1) | | | | |
| School belonging (student, fall year 1) | | | | |
| <i>Missing value indicator variables</i> | | | | |
| English language arts test scores | | + | | + |
| Academic competence | | | | |
| Social Skills Rating System | | | | |
| Altruism | | | | |
| Empathy | | | | |
| Externalizing behavior | | + | | + |
| Aggression | | + | | + |
| Delinquent behavior | | + | | + |
| School expectations (student) | | | | |
| School belonging (student) | | | | |
| <i>Student demographic variables</i> | | | | |
| Grade (student) | + | + | + | + |
| Gender (student) | + | + | + | + |
| English language learner status | + | + | + | + |
| Race/ethnicity | + | + | + | + |
| Parent education | + | + | + | + |
| <i>School-level covariates</i> | | | | |
| Randomization Strata | + | + | + | + |
| California Academic Performance Index scores | + | + | + | + |
| English language arts test scores (grades 2/3, year 0) | + | + | | |
| English language arts test scores (grades 3/4, year 0) | | | + | + |
| Academic competence (grades 2/3, year 0) | | | | |
| Academic competence (grades 3/4, year 0) | | | | |
| Social skills (grades 2/3, year 0) | | | | |
| Social skills (grades 3/4, year 0) | | | | |
| Altruism (grades 4/5, fall year 1) | | | | |
| Empathy (grades 4/5, fall year 1) | | | | |
| Externalizing behavior (grades 2/3, year 0) | + | + | | |
| Externalizing behavior (grades 3/4, year 0) | | | + | + |
| Aggression (grades 4/5, fall year 1) | + | + | + | + |
| Delinquent behavior (grades 4/5, fall year 1) | + | + | + | + |
| School expectations (student, fall year 1) | | | | |
| School belonging (student, fall year 1) | | | | |

Table F4. Covariates included in school environment impact analysis models (students)

| Variable | Primary analysis (year 2) | | Exploratory analysis (year 1) | |
|--|------------------------------|-------------------------------|----------------------------------|-------------------------------|
| | Benchmark model | Missing indicator model | Benchmark model | Missing indicator model |
| <i>Pretest variables</i> | | | | |
| English language arts test scores (spring year 0) | + | + | + | + |
| Academic competence (spring year 0) | | | | |
| Social Skills Rating System (spring year 0) | | | | |
| Altruism (fall year 1) | | | | |
| Empathy (fall year 1) | | | | |
| Externalizing behavior (spring year 0) | | | | |
| Aggression (fall year 1) | | | | |
| Delinquent behavior (fall year 1) | | | | |
| School expectations (student, fall year 1) | | | + | + |
| School belonging (student, fall year 1) | | | + | + |
| <i>Missing value indicator variables</i> | | | | |
| English language arts test scores | | + | | + |
| Academic competence | | | | |
| Social Skills Rating System | | | | |
| Altruism | | | | |
| Empathy | | | | |
| Externalizing behavior | | | | |
| Aggression | | | | |
| Delinquent behavior | | | | |
| School expectations (student) | | | | + |
| School belonging (student) | | | | + |
| <i>Student demographic variables</i> | | | | |
| Grade (student) | + | + | + | + |
| Gender (student) | + | + | + | + |
| English language learner status | + | + | + | + |
| Race/ethnicity | + | + | + | + |
| Parent education | + | + | + | + |
| <i>School-level covariates</i> | | | | |
| Randomization Strata | + | + | + | + |
| California Academic Performance Index scores | + | + | + | + |
| English language arts test scores (grades 2/3, year 0) | + | + | | |
| English language arts test scores (grades 3/4, year 0) | | | + | + |
| Academic competence (grades 2/3, year 0) | | | | |
| Academic competence (grades 3/4, year 0) | | | | |
| Social skills (grades 2/3, year 0) | | | | |
| Social skills (grades 3/4, year 0) | | | | |
| Altruism (grades 4/5, fall year 1) | | | | |
| Empathy (grades 4/5, fall year 1) | | | | |
| Externalizing behavior (grades 2/3, year 0) | | | | |
| Externalizing behavior (grades 3/4, year 0) | | | | |
| Aggression (grades 4/5, fall year 1) | | | | |
| Delinquent behavior (grades 4/5, fall year 1) | | | | |
| School expectations (student, fall year 1) | + | + | + | + |
| School belonging (student, fall year 1) | + | + | + | + |

Table F5. Covariates included in school environment impact analysis models (teachers)

| Variable | Primary analysis (year 2) | | Exploratory analysis (year 1) | |
|--|--------------------------------------|--|--|--|
| | Benchmark model | Missing indicator model | Benchmark model | Missing indicator model |
| <i>Pretest variables</i> | | | | |
| School expectations (teacher, spring year 0) | + | + | + | + |
| School belonging (teacher, spring year 0) | + | + | + | + |
| <i>Missing value indicator variables</i> | | | | |
| School expectations (teacher) | | + | | + |
| School belonging (teacher) | | + | | + |
| <i>Teacher demographic variables</i> | | | | |
| Grade taught | + | + | + | + |
| Gender (student) | + | + | + | + |
| Years of teaching experience | + | + | + | + |
| Race/ethnicity | + | + | + | + |
| <i>School-level covariates</i> | | | | |
| Randomization Strata | + | + | + | + |
| California Academic Performance Index scores | + | + | + | + |
| School expectations (teacher, spring year 0) | + | + | + | + |
| School belonging (teacher, spring year 0) | + | + | + | + |

Appendix G. Sensitivity analyses—primary impacts on student-level outcomes

Table G1. Estimated impact on English language arts test scores across different models, samples, and methods used to handle missing data (primary analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 4) | 0.085 | 0.051 | 1.670 | 0.096 | 0.090 | 4,525 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | −0.001 | 0.068 | −0.010 | 0.990 | −0.001 | 4,525 |
| No covariates except intervention status, pretest, and randomization strata | 0.080 | 0.059 | 1.360 | 0.175 | 0.085 | 4,525 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.024 | 0.048 | 0.500 | 0.619 | 0.025 | 3,875 |
| Included data with missing covariates using missing indicator method | 0.047 | 0.044 | 1.070 | 0.284 | 0.049 | 4,525 |
| Multiple imputation of outcomes and covariates | 0.024 | 0.050 | 0.490 | 0.627 | 0.026 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | 0.076 | 0.053 | 1.440 | 0.151 | 0.080 | 4,271 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 0.090 | 0.052 | 1.720 | 0.085 | 0.095 | 4,468 |
| Student migrants excluded | 0.086 | 0.051 | 1.670 | 0.096 | 0.090 | 4,440 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.050 | 0.043 | 1.160 | 0.246 | 0.053 | 4,525 |

Source: Baseline and year 2 teacher-reported Social Skills Rating System and district archival data.

Table G2. Estimated impact on academic competence across different models, samples, and methods used to handle missing data (primary analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 4) | 0.769 | 0.774 | 0.990 | 0.321 | 0.078 | 4,376 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.027 | 0.493 | -0.060 | 0.956 | -0.003 | 4,376 |
| No covariates except intervention status, pretest, and randomization strata | 0.702 | 0.802 | 0.880 | 0.382 | 0.072 | 4,376 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.245 | 0.487 | 0.500 | 0.615 | 0.025 | 3,782 |
| Included data with missing covariates using missing indicator method | 0.135 | 0.488 | 0.280 | 0.782 | 0.014 | 4,376 |
| Multiple imputation of outcomes and covariates | -0.583 | 0.809 | -0.720 | 0.475 | -0.059 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | 0.770 | 0.814 | 0.950 | 0.344 | 0.078 | 4,166 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 0.962 | 0.572 | 1.680 | 0.094 | 0.098 | 4,333 |
| Student migrants excluded | 0.929 | 0.728 | 1.280 | 0.203 | 0.095 | 4,308 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.193 | 0.491 | 0.390 | 0.695 | 0.020 | 4,376 |

Source: Baseline and year 2 teacher-reported Social Skills Rating System and district archival data.

Table G3. Estimated impact on Social Skills Rating System across different models, samples, and methods used to handle missing data (Primary Analytic Sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 4) | 1.600 | 1.470 | 1.090 | 0.278 | 0.129 | 4,393 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.466 | 1.161 | 0.400 | 0.688 | 0.038 | 4,393 |
| No covariates except intervention status, pretest, and randomization strata | 1.952 | 1.670 | 1.170 | 0.243 | 0.157 | 4,393 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 1.828 | 1.192 | 1.530 | 0.125 | 0.147 | 3,814 |
| Included data with missing covariates using missing indicator method | 0.424 | 1.215 | 0.350 | 0.727 | 0.034 | 4,393 |
| Multiple imputation of outcomes and covariates | -0.063 | 1.624 | -0.040 | 0.969 | -0.005 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | 1.566 | 1.540 | 1.020 | 0.310 | 0.126 | 4,183 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 2.118 | 1.413 | 1.500 | 0.135 | 0.170 | 4,350 |
| Student migrants excluded | 2.335 | 1.453 | 1.610 | 0.109 | 0.188 | 4,325 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.459 | 1.209 | 0.380 | 0.704 | 0.037 | 4,393 |

Source: Baseline teacher-reported Social Skills Rating System and district archival data and year 2 teacher-reported Social Skills Rating System and student survey data.

Table G4. Estimated impact on altruism across different models, samples, and methods used to handle missing data (primary analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 4) | -0.068 | 0.043 | -1.580 | 0.114 | -0.097 | 4,511 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.088 | 0.042 | -2.070 | 0.038 | -0.126 | 4,511 |
| No covariates except intervention status, pretest, and randomization strata | -0.068 | 0.045 | -1.510 | 0.131 | -0.098 | 4,511 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.062 | 0.046 | -1.350 | 0.178 | -0.089 | 3,873 |
| Included data with missing covariates using missing indicator method | -0.083 | 0.042 | -1.980 | 0.047 | -0.120 | 4,511 |
| Multiple imputation of outcomes and covariates | -0.084 | 0.045 | -1.860 | 0.063 | -0.121 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | -0.054 | 0.045 | -1.190 | 0.233 | -0.077 | 4,257 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.063 | 0.044 | -1.430 | 0.153 | -0.090 | 4,457 |
| Student migrants excluded | -0.063 | 0.044 | -1.420 | 0.156 | -0.091 | 4,429 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | -0.088 | 0.045 | -1.960 | 0.050 | -0.127 | 4,511 |

Source: Baseline teacher-reported Social Skills Rating System and district archival data and year 2 teacher-reported Social Skills Rating System and student survey data.

Table G5. Estimated impact on empathy across different models, samples, and methods used to handle missing data (primary analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 4) | -0.013 | 0.024 | -0.540 | 0.588 | -0.034 | 4,512 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.031 | 0.026 | -1.210 | 0.228 | -0.081 | 4,512 |
| No covariates except intervention status, pretest, and randomization strata | -0.019 | 0.027 | -0.710 | 0.475 | -0.050 | 4,512 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.004 | 0.024 | -0.150 | 0.882 | -0.009 | 3,874 |
| Included data with missing covariates using missing indicator method | -0.019 | 0.024 | -0.800 | 0.423 | -0.049 | 4,512 |
| Multiple imputation of outcomes and covariates | -0.021 | 0.025 | -0.850 | 0.396 | -0.055 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | -0.015 | 0.026 | -0.590 | 0.558 | -0.039 | 4,258 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.011 | 0.024 | -0.470 | 0.639 | -0.029 | 4,458 |
| Student migrants excluded | -0.012 | 0.024 | -0.480 | 0.634 | -0.030 | 4,430 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | -0.022 | 0.024 | -0.900 | 0.370 | -0.057 | 4,512 |

Source: Baseline teacher-reported Social Skills Rating System and district archival data and year 2 teacher-reported Social Skills Rating System and student survey data.

Table G6. Estimated impact on externalizing behavior across different models, samples, and methods used to handle missing data (primary analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 4) | -0.392 | 0.351 | -1.120 | 0.268 | -0.153 | 4,379 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.012 | 0.191 | -0.060 | 0.952 | -0.005 | 4,379 |
| No covariates except intervention status, pretest, and randomization strata | -0.503 | 0.448 | -1.120 | 0.265 | -0.196 | 4,379 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.219 | 0.189 | -1.160 | 0.247 | -0.085 | 3,789 |
| Included data with missing covariates using missing indicator method | -0.164 | 0.175 | -0.940 | 0.350 | -0.064 | 4,379 |
| Multiple imputation of outcomes and covariates | -0.162 | 0.365 | -0.440 | 0.660 | -0.063 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | -0.405 | 0.362 | -1.120 | 0.266 | -0.158 | 4,169 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.413 | 0.358 | -1.160 | 0.251 | -0.161 | 4,336 |
| Student migrants excluded | -0.363 | 0.341 | -1.070 | 0.289 | -0.141 | 4,311 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | -0.161 | 0.177 | -0.910 | 0.362 | -0.063 | 4,379 |

Source: Baseline teacher-reported Social Skills Rating System and district archival data and year 2 teacher-reported Social Skills Rating System and student survey data.

Table G7. Estimated impact on aggression across different models, samples, and methods used to handle missing data (primary analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 4) | 0.023 | 0.045 | 0.510 | 0.609 | 0.044 | 4,507 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.049 | 0.040 | 1.220 | 0.224 | 0.094 | 4,507 |
| No covariates except intervention status, pretest, and randomization strata | 0.019 | 0.047 | 0.410 | 0.681 | 0.037 | 4,507 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.044 | 0.043 | 1.020 | 0.306 | 0.085 | 3,854 |
| Included data with missing covariates using missing indicator method | 0.036 | 0.041 | 0.880 | 0.380 | 0.071 | 4,507 |
| Multiple imputation of outcomes and covariates | 0.031 | 0.046 | 0.670 | 0.501 | 0.060 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | 0.026 | 0.046 | 0.560 | 0.572 | 0.050 | 4,254 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 0.017 | 0.044 | 0.390 | 0.699 | 0.033 | 4,453 |
| Student migrants excluded | 0.023 | 0.044 | 0.530 | 0.595 | 0.045 | 4,425 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.041 | 0.042 | 0.960 | 0.335 | 0.079 | 4,507 |

Source: Baseline teacher-reported Social Skills Rating System and district archival data and year 2 teacher-reported Social Skills Rating System and student survey data.

Table G8. Estimated impact on delinquent behavior across different models, samples, and methods used to handle missing data (primary analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 4) | 0.000 | 0.026 | 0.000 | 0.996 | 0.000 | 4,507 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.015 | 0.023 | 0.640 | 0.520 | 0.047 | 4,507 |
| No covariates except intervention status, pretest, and randomization strata | -0.004 | 0.028 | -0.130 | 0.899 | -0.011 | 4,507 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.017 | 0.026 | 0.650 | 0.515 | 0.053 | 3,854 |
| Included data with missing covariates using missing indicator method | 0.008 | 0.024 | 0.340 | 0.736 | 0.025 | 4,507 |
| Multiple imputation of outcomes and covariates | 0.006 | 0.026 | 0.230 | 0.820 | 0.019 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | -0.001 | 0.027 | -0.040 | 0.967 | -0.004 | 4,254 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.006 | 0.026 | -0.240 | 0.812 | -0.020 | 4,453 |
| Student migrants excluded | -0.004 | 0.026 | -0.170 | 0.866 | -0.014 | 4,425 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.012 | 0.025 | 0.470 | 0.641 | 0.037 | 4,507 |

Source: Baseline teacher-reported Social Skills Rating System and district archival data and year 2 teacher-reported Social Skills Rating System and student survey data.

Table G9. Estimated impact on school expectations across different models, samples, and methods used to handle missing data (primary analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 4) | 0.011 | 0.063 | 0.170 | 0.864 | 0.013 | 4,511 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.008 | 0.067 | -0.120 | 0.902 | -0.010 | 4,511 |
| No covariates except intervention status, pretest, and randomization strata | -0.002 | 0.066 | -0.030 | 0.974 | -0.003 | 4,511 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.016 | 0.064 | 0.250 | 0.803 | 0.019 | 4,160 |
| Included data with missing covariates using missing indicator method | 0.010 | 0.063 | 0.160 | 0.876 | 0.012 | 4,511 |
| Multiple imputation of outcomes and covariates | 0.007 | 0.063 | 0.110 | 0.910 | 0.009 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | 0.008 | 0.066 | 0.120 | 0.903 | 0.010 | 4,257 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 0.012 | 0.064 | 0.190 | 0.848 | 0.015 | 4,457 |
| Student migrants excluded | 0.014 | 0.065 | 0.210 | 0.833 | 0.016 | 4,429 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.008 | 0.063 | 0.120 | 0.902 | 0.009 | 4,511 |

Source: Baseline student survey (aggregated by school) and district archival data and year 2 student survey data.

Table G10. Estimated impact on student feelings of belonging across different models, samples, and methods used to handle missing data (primary analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 4) | -0.031 | 0.066 | -0.480 | 0.632 | -0.044 | 4,514 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.070 | 0.074 | -0.950 | 0.342 | -0.099 | 4,514 |
| No covariates except intervention status, pretest, and randomization strata | -0.072 | 0.074 | -0.980 | 0.329 | -0.102 | 4,514 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.039 | 0.064 | -0.610 | 0.544 | -0.054 | 4,163 |
| Included data with missing covariates using missing indicator method | -0.031 | 0.066 | -0.470 | 0.636 | -0.044 | 4,514 |
| Multiple imputation of outcomes and covariates | -0.031 | 0.066 | -0.470 | 0.637 | -0.044 | 4,683 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | -0.036 | 0.070 | -0.520 | 0.602 | -0.051 | 4,260 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.031 | 0.065 | -0.480 | 0.633 | -0.044 | 4,460 |
| Student migrants excluded | -0.040 | 0.064 | -0.620 | 0.536 | -0.056 | 4,432 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | -0.039 | 0.065 | -0.600 | 0.552 | -0.055 | 4,514 |

Source: Baseline student survey (aggregated by school) and district archival data and year 2 student survey data.

Appendix H. Sensitivity analyses—primary impacts on teacher-level outcomes

Table H1. Estimated impact on teacher reports of school expectations across different models, samples, and methods used to handle missing data (primary analytic sample – grades 2-5)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 4) | 0.034 | 0.061 | 0.550 | 0.583 | 0.070 | 507 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.008 | 0.055 | 0.150 | 0.880 | 0.017 | 507 |
| No covariates except intervention status, pretest, and randomization strata | 0.028 | 0.059 | 0.480 | 0.632 | 0.059 | 507 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.001 | 0.054 | -0.010 | 0.988 | -0.002 | 427 |
| Included data with missing covariates using missing indicator method | 0.022 | 0.051 | 0.440 | 0.661 | 0.047 | 507 |
| Multiple imputation of outcomes and covariates | 0.028 | 0.059 | 0.480 | 0.631 | 0.059 | 550 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | 0.032 | 0.063 | 0.510 | 0.611 | 0.069 | 467 |

I-H

Source: Baseline and year 2 teacher survey data.

Table H2. Estimated impact on teacher reports of student feelings of belonging across different models, samples, and methods used to handle missing data (primary analytic sample – grades 2-5)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 4) | 0.074 | 0.070 | 1.060 | 0.289 | 0.130 | 507 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.010 | 0.082 | 0.120 | 0.906 | 0.017 | 507 |
| No covariates except intervention status, pretest, and randomization strata | 0.038 | 0.077 | 0.490 | 0.623 | 0.067 | 507 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.014 | 0.068 | 0.200 | 0.840 | 0.024 | 427 |
| Included data with missing covariates using missing indicator method | 0.022 | 0.066 | 0.330 | 0.740 | 0.039 | 507 |
| Multiple imputation of outcomes and covariates | 0.074 | 0.072 | 1.030 | 0.302 | 0.130 | 550 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | 0.081 | 0.073 | 1.110 | 0.269 | 0.153 | 467 |

Source: Baseline and year 2 teacher survey data.

Table H3. Estimated impact on teacher reports of school expectations across different models, samples, and methods used to handle missing data (primary analytic sample – grades 4-5)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 4) | 0.008 | 0.084 | 0.100 | 0.923 | 0.016 | 209 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.029 | 0.093 | 0.310 | 0.754 | 0.058 | 209 |
| No covariates except intervention status, pretest, and randomization strata | -0.002 | 0.098 | -0.020 | 0.980 | -0.005 | 209 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.030 | 0.094 | 0.320 | 0.752 | 0.059 | 174 |
| Included data with missing covariates using missing indicator method | 0.063 | 0.085 | 0.740 | 0.460 | 0.125 | 209 |
| Multiple imputation of outcomes and covariates | 0.196 | 0.298 | 0.660 | 0.516 | 0.393 | 229 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | 0.042 | 0.087 | 0.490 | 0.624 | 0.089 | 192 |

Source: Baseline and year 2 teacher survey data.

Table H4. Estimated impact on teacher reports of student feelings of belonging across different models, samples, and methods used to handle missing data (primary analytic sample – grades 4-5)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 4) | 0.051 | 0.101 | 0.500 | 0.617 | 0.081 | 209 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.036 | 0.111 | 0.320 | 0.749 | 0.056 | 209 |
| No covariates except intervention status, pretest, and randomization strata | 0.051 | 0.110 | 0.460 | 0.645 | 0.081 | 209 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.009 | 0.088 | 0.100 | 0.918 | 0.014 | 174 |
| Included data with missing covariates using missing indicator method | 0.054 | 0.094 | 0.570 | 0.569 | 0.086 | 209 |
| Multiple imputation of outcomes and covariates | 0.326 | 0.442 | 0.740 | 0.465 | 0.518 | 229 |
| <i>School attrition</i> | | | | | | |
| Excluded randomization strata in which schools dropped out of study | 0.079 | 0.105 | 0.760 | 0.450 | 0.130 | 192 |

Source: Baseline and year 2 teacher survey data.

Appendix I. Sensitivity analyses—exploratory impacts on student-level outcomes

Table I1. Estimated impact on English language arts test scores across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| <i>Benchmark model (chapter 5)</i> | -0.018 | 0.043 | -0.420 | 0.674 | -0.019 | 5,355 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.009 | 0.068 | 0.140 | 0.891 | 0.010 | 5,355 |
| No covariates except intervention status, pretest, and randomization strata | -0.014 | 0.045 | -0.320 | 0.753 | -0.015 | 5,355 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.019 | 0.038 | -0.500 | 0.620 | -0.020 | 4,528 |
| Included data with missing covariates using missing indicator method | -0.010 | 0.035 | -0.280 | 0.780 | -0.011 | 5,355 |
| Multiple imputation of outcomes and covariates | 0.033 | 0.174 | 0.190 | 0.849 | 0.036 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.022 | 0.044 | -0.510 | 0.614 | -0.024 | 5,301 |
| Student migrants excluded | -0.023 | 0.044 | -0.520 | 0.602 | -0.025 | 5,282 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | -0.010 | 0.035 | -0.290 | 0.775 | -0.011 | 5,355 |

Source: Baseline and year 1 teacher-reported Social Skills Rating System and district archival data.

Table I2. Estimated impact on academic competence across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 5) | 0.130 | 1.021 | 0.130 | 0.899 | 0.013 | 5,120 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.377 | 0.441 | 0.860 | 0.392 | 0.038 | 5,120 |
| No covariates except intervention status, pretest, and randomization strata | -0.173 | 0.998 | -0.170 | 0.863 | -0.017 | 5,120 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.449 | 0.426 | 1.050 | 0.293 | 0.045 | 4,290 |
| Included data with missing covariates using missing indicator method | 0.457 | 0.380 | 1.200 | 0.229 | 0.046 | 5,120 |
| Multiple imputation of outcomes and covariates | -1.059 | 3.207 | -0.330 | 0.741 | -0.105 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 0.079 | 1.046 | 0.080 | 0.940 | 0.008 | 5,093 |
| Student migrants excluded | 0.107 | 1.059 | 0.100 | 0.920 | 0.011 | 5,082 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.471 | 0.378 | 1.250 | 0.213 | 0.047 | 5,120 |

I
2

Source: Baseline and year 1 teacher-reported Social Skills Rating System and district archival data.

Table I3. Estimated impact on Social Skills Rating System across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 5) | 0.702 | 3.070 | 0.230 | 0.820 | 0.053 | 5,149 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.047 | 0.966 | 0.050 | 0.961 | 0.004 | 5,149 |
| No covariates except intervention status, pretest, and randomization strata | 0.025 | 2.625 | 0.010 | 0.992 | 0.002 | 5,149 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 1.114 | 1.049 | 1.060 | 0.288 | 0.084 | 4,155 |
| Included data with missing covariates using missing indicator method | 1.352 | 0.949 | 1.430 | 0.154 | 0.102 | 5,149 |
| Multiple imputation of outcomes and covariates | 1.388 | 7.055 | 0.200 | 0.844 | 0.105 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 0.471 | 3.158 | 0.150 | 0.882 | 0.036 | 5,122 |
| Student migrants excluded | 0.568 | 3.180 | 0.180 | 0.859 | 0.043 | 5,111 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 1.381 | 0.936 | 1.470 | 0.140 | 0.105 | 5,149 |

1

Source: Baseline teacher-reported Social Skills Rating System and district archival data and year 1 teacher-reported Social Skills Rating System and student survey data.

Table I4. Estimated impact on altruism across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 5) | -0.073 | 0.031 | -2.400 | 0.016 | -0.107 | 5,410 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.077 | 0.036 | -2.140 | 0.032 | -0.113 | 5,410 |
| No covariates except intervention status, pretest, and randomization strata | -0.062 | 0.031 | -1.980 | 0.048 | -0.090 | 5,410 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.062 | 0.032 | -1.960 | 0.050 | -0.091 | 4,332 |
| Included data with missing covariates using missing indicator method | -0.079 | 0.029 | -2.730 | 0.006 | -0.115 | 5,410 |
| Multiple imputation of outcomes and covariates | -0.070 | 0.031 | -2.220 | 0.026 | -0.102 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.075 | 0.031 | -2.440 | 0.015 | -0.111 | 5,361 |
| Student migrants excluded | -0.074 | 0.030 | -2.430 | 0.015 | -0.108 | 5,344 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | -0.079 | 0.030 | -2.690 | 0.007 | -0.116 | 5,410 |

I
4

Source: Baseline teacher-reported Social Skills Rating System and district archival data and year 1 teacher-reported Social Skills Rating System and student survey data.

Table I5. Estimated impact on empathy across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 5) | -0.033 | 0.020 | -1.640 | 0.101 | -0.089 | 5,410 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.029 | 0.022 | -1.300 | 0.194 | -0.077 | 5,410 |
| No covariates except intervention status, pretest, and randomization strata | -0.027 | 0.020 | -1.400 | 0.162 | -0.073 | 5,410 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.034 | 0.019 | -1.750 | 0.080 | -0.090 | 4,333 |
| Included data with missing covariates using missing indicator method | -0.036 | 0.019 | -1.900 | 0.058 | -0.097 | 5,410 |
| Multiple imputation of outcomes and covariates | -0.033 | 0.020 | -1.620 | 0.106 | -0.088 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.034 | 0.021 | -1.620 | 0.106 | -0.089 | 5,361 |
| Student migrants excluded | -0.032 | 0.021 | -1.520 | 0.130 | -0.084 | 5,344 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | -0.037 | 0.021 | -1.780 | 0.076 | -0.097 | 5,410 |

51

Source: Baseline teacher-reported Social Skills Rating System and district archival data and year 1 teacher-reported Social Skills Rating System and student survey data.

Table I6. Estimated impact on externalizing behavior across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 5) | -0.321 | 0.846 | -0.380 | 0.706 | -0.116 | 5,120 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.025 | 0.172 | 0.150 | 0.884 | 0.009 | 5,120 |
| No covariates except intervention status, pretest, and randomization strata | -0.341 | 0.834 | -0.410 | 0.683 | -0.123 | 5,120 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.049 | 0.162 | -0.300 | 0.761 | -0.018 | 4,093 |
| Included data with missing covariates using missing indicator method | -0.101 | 0.149 | -0.680 | 0.497 | -0.037 | 5,120 |
| Multiple imputation of outcomes and covariates | 0.986 | 2.501 | 0.390 | 0.694 | 0.356 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.295 | 0.843 | -0.350 | 0.727 | -0.107 | 5,093 |
| Student migrants excluded | -0.299 | 0.843 | -0.350 | 0.724 | -0.108 | 5,082 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | -0.101 | 0.150 | -0.670 | 0.501 | -0.036 | 5,120 |

Table I7. Estimated impact on aggression across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 5) | 0.023 | 0.033 | 0.710 | 0.480 | 0.043 | 5,390 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.035 | 0.029 | 1.190 | 0.233 | 0.066 | 5,390 |
| No covariates except intervention status, pretest, and randomization strata | 0.024 | 0.032 | 0.740 | 0.462 | 0.045 | 5,390 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.033 | 0.028 | 1.210 | 0.228 | 0.062 | 4,276 |
| Included data with missing covariates using missing indicator method | 0.033 | 0.026 | 1.270 | 0.205 | 0.062 | 5,390 |
| Multiple imputation of outcomes and covariates | 0.028 | 0.031 | 0.890 | 0.376 | 0.053 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 0.027 | 0.033 | 0.810 | 0.420 | 0.050 | 5,341 |
| Student migrants excluded | 0.026 | 0.032 | 0.810 | 0.418 | 0.049 | 5,324 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.033 | 0.026 | 1.240 | 0.216 | 0.062 | 5,390 |

71

Source: Baseline and year 1 teacher-reported Social Skills Rating System, district archival, and student survey data.

Table I8. Estimated impact on delinquent behavior across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 5) | 0.013 | 0.019 | 0.700 | 0.483 | 0.042 | 5,391 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.015 | 0.018 | 0.830 | 0.405 | 0.048 | 5,391 |
| No covariates except intervention status, pretest, and randomization strata | 0.017 | 0.019 | 0.850 | 0.397 | 0.052 | 5,391 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.021 | 0.018 | 1.170 | 0.242 | 0.067 | 4,275 |
| Included data with missing covariates using missing indicator method | 0.015 | 0.016 | 0.920 | 0.358 | 0.048 | 5,391 |
| Multiple imputation of outcomes and covariates | 0.016 | 0.019 | 0.860 | 0.390 | 0.052 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 0.016 | 0.020 | 0.770 | 0.439 | 0.049 | 5,342 |
| Student migrants excluded | 0.014 | 0.020 | 0.720 | 0.472 | 0.046 | 5,325 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.015 | 0.017 | 0.910 | 0.363 | 0.048 | 5,391 |

Table I9. Estimated impact on school expectations across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 5) | -0.006 | 0.046 | -0.140 | 0.891 | -0.007 | 5,413 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.002 | 0.056 | -0.040 | 0.969 | -0.002 | 5,413 |
| No covariates except intervention status, pretest, and randomization strata | -0.014 | 0.046 | -0.290 | 0.771 | -0.016 | 5,413 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.016 | 0.049 | 0.320 | 0.748 | 0.018 | 4,668 |
| Included data with missing covariates using missing indicator method | 0.010 | 0.044 | 0.240 | 0.813 | 0.012 | 5,413 |
| Multiple imputation of outcomes and covariates | 0.000 | 0.052 | 0.010 | 0.995 | 0.000 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | -0.013 | 0.045 | -0.280 | 0.780 | -0.015 | 5,364 |
| Student migrants excluded | -0.011 | 0.046 | -0.230 | 0.816 | -0.012 | 5,347 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.012 | 0.043 | 0.270 | 0.788 | 0.013 | 5,413 |

Table I10. Estimated impact on student feelings of belonging across different models, samples, and methods used to handle missing data (exploratory analytic sample)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|-----------------|----------------|-------------|---------|-------------|-------------|
| Benchmark model (chapter 5) | 0.010 | 0.056 | 0.180 | 0.860 | 0.014 | 5,415 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | -0.027 | 0.067 | -0.410 | 0.685 | -0.038 | 5,415 |
| No covariates except intervention status, pretest, and randomization strata | -0.007 | 0.055 | -0.130 | 0.898 | -0.010 | 5,415 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | -0.007 | 0.053 | -0.130 | 0.898 | -0.009 | 4,670 |
| Included data with missing covariates using missing indicator method | 0.001 | 0.052 | 0.020 | 0.980 | 0.002 | 5,415 |
| Multiple imputation of outcomes and covariates | 0.012 | 0.057 | 0.200 | 0.839 | 0.016 | 5,674 |
| <i>Crossovers and migrants</i> | | | | | | |
| Student crossovers excluded | 0.007 | 0.057 | 0.130 | 0.898 | 0.010 | 5,366 |
| Student migrants excluded | 0.010 | 0.057 | 0.170 | 0.864 | 0.014 | 5,349 |
| <i>Weighting</i> | | | | | | |
| Nonresponse weights in model based on missing indicator method | 0.002 | 0.051 | 0.030 | 0.975 | 0.002 | 5,415 |

Appendix J. Sensitivity analyses—exploratory impacts on teacher-level outcomes

Table J1. Estimated impact on teacher reports of school expectations across different models, samples, and methods used to handle missing data (exploratory analytic sample – grades 2-5)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 5) | 0.076 | 0.043 | 1.760 | 0.079 | 0.156 | 564 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.073 | 0.048 | 1.510 | 0.132 | 0.149 | 564 |
| No covariates except intervention status, pretest, and randomization strata | 0.065 | 0.045 | 1.450 | 0.148 | 0.134 | 564 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.078 | 0.042 | 1.870 | 0.061 | 0.158 | 508 |
| Included data with missing covariates using missing indicator method | 0.089 | 0.039 | 2.280 | 0.023 | 0.181 | 564 |
| Multiple imputation of outcomes and covariates | 0.056 | 0.063 | 0.880 | 0.381 | 0.115 | 625 |

Source: Baseline and year 1 teacher survey data.

Table J2. Estimated impact on teacher reports of student feelings of belonging across different models, samples, and methods used to handle missing data (exploratory analytic sample – grades 2-5)

| Method | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 5) | 0.061 | 0.064 | 0.950 | 0.343 | 0.106 | 564 |
| <i>Covariate adjustment</i> | | | | | | |
| No covariates except intervention status and randomization strata | 0.071 | 0.082 | 0.870 | 0.383 | 0.124 | 564 |
| No covariates except intervention status, pretest, and randomization strata | 0.055 | 0.066 | 0.840 | 0.401 | 0.096 | 564 |
| <i>Missing data adjustment</i> | | | | | | |
| Excluded cases with any missing data (listwise deletion) | 0.055 | 0.058 | 0.940 | 0.345 | 0.099 | 508 |
| Included data with missing covariates using missing indicator method | 0.083 | 0.057 | 1.440 | 0.150 | 0.143 | 564 |
| Multiple imputation of outcomes and covariates | 0.079 | 0.072 | 1.100 | 0.273 | 0.137 | 625 |

Source: Baseline and year 1 teacher survey data.

Table J3. Estimated impact on teacher reports of school expectations across different models, samples, and methods used to handle missing data (exploratory analytic sample – grades 4-5)

| Method | | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|--|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 5) | | 0.114 | 0.076 | 1.50 | 0.134 | 0.221 | 210 |
| <i>Covariate adjustment</i> | | | | | | | |
| No covariates except intervention status and randomization strata | | 0.010 | 0.081 | 0.12 | 0.903 | 0.019 | 210 |
| No covariates except intervention status, pretest, and randomization strata | | 0.069 | 0.083 | 0.83 | 0.409 | 0.133 | 210 |
| <i>Missing data adjustment</i> | | | | | | | |
| Excluded cases with any missing data (listwise deletion) | | 0.102 | 0.079 | 1.28 | 0.199 | 0.195 | 192 |
| Included data with missing covariates using missing indicator method | | 0.097 | 0.075 | 1.29 | 0.198 | 0.186 | 210 |
| Multiple imputation of outcomes and covariates | | 1.054 | 0.706 | 1.49 | 0.138 | 2.032 | 240 |

Source: Baseline and year 1 teacher survey data.

Table J4. Estimated impact on teacher reports of student feelings of belonging across different models, samples, and methods used to handle missing data (exploratory analytic sample – grades 4-5)

| Method | | Impact estimate | Standard error | z-statistic | p-value | Effect size | Sample size |
|---|--|------------------------|-----------------------|--------------------|----------------|--------------------|--------------------|
| Benchmark model (chapter 5) | | 0.068 | 0.102 | 0.67 | 0.506 | 0.112 | 210 |
| <i>Covariate adjustment</i> | | | | | | | |
| No covariates except intervention status and randomization strata | | 0.005 | 0.112 | 0.04 | 0.964 | 0.008 | 210 |
| No covariates except intervention status, pretest, and randomization strata | | 0.032 | 0.115 | 0.27 | 0.784 | 0.052 | 210 |
| <i>Missing data adjustment</i> | | | | | | | |
| Excluded cases with any missing data (listwise deletion) | | 0.035 | 0.089 | 0.39 | 0.699 | 0.060 | 192 |
| Included data with missing covariates using missing indicator method | | 0.055 | 0.092 | 0.60 | 0.551 | 0.091 | 210 |
| Multiple imputation of outcomes and covariates | | -0.279 | 1.072 | -0.26 | 0.795 | -0.460 | 240 |

Source: Baseline and year 1 teacher survey data.

Appendix K. Teacher survey

Expiration Date: 5/31/2010

Teacher Report

Background, School Climate, Practices, and Activities

NOTICE: According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this information collection is 1850-0826. The time required to complete this information collection is estimated to average 15 minutes per student (approximately 3 hours total). If you have any comments concerning the accuracy of the time estimate(s) or suggestions for improving this form, please write to: U.S. Department of Education, Washington, D.C. 20202-4651. If you have comments or concerns regarding the status of your individual submission of this form, write directly to: Rafael Valdivieso, U.S. Department of Education, 555 New Jersey Avenue, NW, Room 506E, Washington, D.C. 20208.

Responses to this data collection will be used only for statistical purposes. The reports prepared for this study will summarize findings across the sample and will not associate responses with a specific district or individual. We will not provide information that identifies you or your school or your district to anyone outside the study team, except as required by law.

LESSONS IN CHARACTER STUDY

TEACHER REPORT

Grades 2-5

Thank you, in advance, for responding to the following questions. This instrument is being administered as part of a long-term research project by WestEd to examine the impact of the *Lessons in Character Program* (LIC) — an English language arts-based character education program — on student academic performance, attendance, school motivation, and endorsement of universal values consistent with character education. WestEd, a research, development, and service agency, works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults.

This survey is designed for you to provide us with background information and to share with us your perceptions about students and staff in your school.

Please read each question and bubble in your response.

Like this: Not like this:

Example:

| | Strongly Disagree | Disagree | Don't Know | Agree | Strongly Agree |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| I enjoy teaching Math | <input type="radio"/> |

DEMOGRAPHIC BACKGROUND AND PROFESSIONAL DEVELOPMENT EXPERIENCES

1. Are you male or female?
 Male
 Female
2. Are you of Hispanic or Latino origin? (Please select only one)
 Yes
 No
3. What is your race? (Please select one or more)
 American Indian or Alaska Native
 Asian
 Black or African American
 Native Hawaiian or Other Pacific Islander
 White
4. Including this year, how many years have you been employed as a teacher?

Years

5. Including this year, how many years have you been employed as a teacher at this school?

Years

6. Considering all of the professional development activities in which you participated in the last 12 months (excluding preservice training), how many total hours, if any, have you spent in activities in which the following content areas were a major focus?

Total hours spent in last 12 months

| 0 | 1-8 | 9-32 | 32+ |
|----------|------------|-------------|------------|
|----------|------------|-------------|------------|

| | | | | |
|--|-----------------------|-----------------------|-----------------------|-----------------------|
| a. New methods of teaching (such as, cooperative learning) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| b. Student performance assessment (such as, methods of testing, applying results to modify instruction). | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| c. Classroom management, including student discipline. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

STUDENTS IN THIS SCHOOL

Mark your agreement or disagreement with the following.

| The students in this school... | | | | | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | Strongly Disagree | Disagree | Don't Know | Agree | Strongly Agree |
| 7. Are nice to each other. | <input type="radio"/> |
| 8. Work together on things. | <input type="radio"/> |
| 9. Make new students feel welcome. | <input type="radio"/> |
| 10. Get along well together. | <input type="radio"/> |
| 11. Work out problems without fighting. | <input type="radio"/> |
| 12. Take good care of school property. | <input type="radio"/> |
| 13. Take responsibility for what they do. | <input type="radio"/> |
| 14. Respect their teachers. | <input type="radio"/> |
| 15. Think it's important to be a good citizen. | <input type="radio"/> |
| 16. Treat one another fairly. | <input type="radio"/> |
| 17. Tell the truth. | <input type="radio"/> |
| 18. Are treated fairly by the adults in the school. | <input type="radio"/> |
| 19. Are expected to get along. | <input type="radio"/> |
| 20. Are expected to be kind and caring. | <input type="radio"/> |
| 21. Are expected to treat each other fairly. | <input type="radio"/> |
| 22. Are expected to obey the rules. | <input type="radio"/> |
| 23. Are expected to tell the truth. | <input type="radio"/> |

SCHOOL STAFF

Mark your agreement or disagreement with the following.

| | Strongly Disagree | Disagree | Don't Know | Agree | Strongly Agree |
|--|-------------------|----------|------------|-------|----------------|
|--|-------------------|----------|------------|-------|----------------|

The school staff...

| | | | | | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 24. Cares about the students' families. | <input type="radio"/> |
| 25. Treats parents with respect. | <input type="radio"/> |
| 26. Makes parents feel welcome at school. | <input type="radio"/> |
| 27. Values parents' ideas and input. | <input type="radio"/> |
| 28. Encourages parents to be involved at school. | <input type="radio"/> |
| 29. Communicates effectively with parents. | <input type="radio"/> |
| 30. Tells parents about the school's efforts to develop good character in students. | <input type="radio"/> |
| 31. Makes conscious effort to develop students' character. | <input type="radio"/> |
| 32. Falls into conflicting cliques. | <input type="radio"/> |
| 33. Is supportive of one another. | <input type="radio"/> |
| 34. Can be counted on to help each other out. | <input type="radio"/> |
| 35. Shares the same beliefs about what the central mission of the school should be. | <input type="radio"/> |
| 36. Frequently consults with and helps one another. | <input type="radio"/> |
| 37. Cooperates with one another. | <input type="radio"/> |
| 38. Seems like a big family with everyone close and cordial. | <input type="radio"/> |
| 39. Supports me when I try out new ideas. | <input type="radio"/> |
| 40. Provides good counsel when I have a teaching problem. | <input type="radio"/> |

THIS SCHOOL

Mark your agreement or disagreement with the following.

| | Strongly Disagree | Disagree | Don't Know | Agree | Strongly Agree |
|--|-------------------|----------|------------|-------|----------------|
|--|-------------------|----------|------------|-------|----------------|

In this school...

| | | | | | |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 41. Both administrators and teachers take active roles in school activities. | <input type="radio"/> |
| 42. We have the needed resources to get the job done. | <input type="radio"/> |
| 43. Things are well organized. | <input type="radio"/> |
| 44. Staff is recognized for a job well done. | <input type="radio"/> |
| 45. Staff is involved in decisions that affect them. | <input type="radio"/> |
| 46. There is interest in innovation and new ideas. | <input type="radio"/> |

Teachers' approaches to classroom management and instruction can differ greatly within a school and between schools. The statements below reflect different approaches teachers might implement in their classrooms. Please rate the extent to which each approach is true in your classroom.

Mark your agreement or disagreement with the following.

| | Never | Sometimes | Often | Almost Always | Always |
|--|-------|-----------|-------|---------------|--------|
|--|-------|-----------|-------|---------------|--------|

In my classroom...

| | | | | | |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 47. Students play an active role in classroom governance or rule-setting. | <input type="radio"/> |
| 48. When problems arise, class meetings or group problem-solving strategies are used. | <input type="radio"/> |
| 49. Social and character development themes are integrated into academic topics (such as historical instruction, literature containing particular values, etc.). | <input type="radio"/> |
| 50. Instructional strategies tend to be student-centered rather than teacher driven. | <input type="radio"/> |
| 51. Discipline strategies focus on social and character development rather than punishment. | <input type="radio"/> |

ACTIVITIES YOU HAVE IMPLEMENTED IN YOUR CLASSROOM

52. For each of the following, please indicate in column 1 whether you have **implemented the activity in your classroom in the last 6 months.**

If so, write the formal name and description of the activity in column 2. If the activity has no formal name, just describe it.

Finally, we would also like to know the frequency and duration of the activity or program in column 3 (for instance, whether it was daily for 6 weeks or 10 sessions over 10 weeks.)?

| a. Character education (promoting positive character traits such as honesty, respect, and cooperation) | <input type="radio"/> <input type="radio"/> <input type="radio"/> | None/Not implemented Some students All students | | |
|---|---|---|--|--|
| b. Violence prevention (changing attitudes, values and/or behaviors regarding violence and aggression) | <input type="radio"/> <input type="radio"/> <input type="radio"/> | None/Not implemented Some students All students | | |
| c. Tolerance, diversity, cultural awareness, or bias awareness (understanding cultural differences, reducing/preventing prejudice or strained racial relations) | <input type="radio"/> <input type="radio"/> <input type="radio"/> | None/Not implemented Some students All students | | |

ACTIVITIES YOU HAVE IMPLEMENTED IN YOUR CLASSROOM

| | | (1) Implemented among which students? | (2) Name and/or description of program activity | (3) Duration or frequency of program activity |
|---|---|---|--|---|
| d. Civics or citizenship (fostering a sense of belonging and a sense of responsibility to the community) | <input type="radio"/> <input type="radio"/> <input type="radio"/> | None/Not implemented Some students All students | | |
| e. Community service or service learning (promoting helping behaviors, "giving back" to the community) | <input type="radio"/> <input type="radio"/> <input type="radio"/> | None/Not implemented Some students All students | | |
| f. Targeted risk prevention (changing behaviors related to current or future risk behaviors such as drug/alcohol use) | <input type="radio"/> <input type="radio"/> <input type="radio"/> | None/Not implemented Some students All students | | |
| g. Self-management/self-discipline (increasing students' knowledge of and control over their own behavioral choices) | <input type="radio"/> <input type="radio"/> <input type="radio"/> | None/Not implemented Some students All students | | |
| h. Emotional control/anger management (such as managing of intense emotions, relaxation, self-calming techniques) | <input type="radio"/> <input type="radio"/> <input type="radio"/> | None/Not implemented Some students All students | | |

ACTIVITIES YOU HAVE IMPLEMENTED IN YOUR CLASSROOM

| | | (1) Implemented among which students? | (2) Name and/or description of program activity | (3) Duration or frequency of program activity |
|---|---|---|--|---|
| i. Resisting peer pressure (includes refusal skills, avoiding risky situations) | <input type="radio"/> None/Not implemented <input type="radio"/> Some students <input type="radio"/> All students | | | |
| j. Perspective-taking/empathy (understanding others' points of view, understanding how one's behavior affects others) | <input type="radio"/> None/Not implemented <input type="radio"/> Some students <input type="radio"/> All students | | | |
| k. Assertiveness/communication skills (includes listening skills, successful communication and/or negotiation strategies) | <input type="radio"/> None/Not implemented <input type="radio"/> Some students <input type="radio"/> All students | | | |
| l. Conflict resolution/social problem solving (such as managing or avoiding conflict, seeking out a third party to mediate) | <input type="radio"/> None/Not implemented <input type="radio"/> Some students <input type="radio"/> All students | | | |
| m. Individual behavior management (such as daily/weekly behavior checklists or report cards, individual point or token reward system) | <input type="radio"/> None/Not implemented <input type="radio"/> Some students <input type="radio"/> All students | | | |

ACTIVITIES YOU HAVE IMPLEMENTED IN YOUR CLASSROOM

| | | (1) Implemented among which students? | (2) Name and/or description of program activity | (3) Duration or frequency of program activity |
|--|---|---|--|---|
| n. Group behavior management (such as daily/weekly behavior checklists or report cards, individual point or token reward system) | <input type="radio"/> None/Not implemented <input type="radio"/> Some students <input type="radio"/> All students | | | |
| o. Other activity #1 (please specify) | <input type="radio"/> None/Not implemented <input type="radio"/> Some students <input type="radio"/> All students | | | |
| p. Other activity #2 (please specify) | <input type="radio"/> None/Not implemented <input type="radio"/> Some students <input type="radio"/> All students | | | |

53. What types of materials (if any) have you used in conjunction with the above activities (#52)?
Mark all that apply.

- Teacher/Instructor guides (curriculum, manual, etc.)
- Materials for student completion (workbooks, worksheets, etc.)
- Instructional aids (games, computer software, videos, puppets, etc.)
- Giveaways (bookmarks, stickers, etc.)
- Other (please specify)
- NONE OF THE ABOVE

54. Some teachers work independently while other teachers prefer to get input from other teachers. Would you say you get...

- No input
- Minimal input
- Moderate input
- A great deal of input

55. How comfortable are you *receiving* advice from other teachers?

- Not at all comfortable
- Slightly comfortable
- Moderately comfortable
- Completely comfortable

55. How comfortable are you *receiving* advice from other teachers?

- Not at all comfortable
- Slightly comfortable
- Moderately comfortable
- Completely comfortable

57. How supportive are other teachers at your school when you need help or advice with teaching?

- Virtually no teachers are supportive
- Some teachers are supportive, but a majority are not
- A majority of teachers are supportive, but some are not
- Nearly every teacher is supportive

58. How receptive are other teachers at your school when you offer help or advice with teaching?

- Virtually no teachers are receptive
- Some teachers are receptive, but a majority are not
- A majority of teachers are receptive, but some are not
- Nearly every teacher is receptive

59. Think of changes that you have made ***over the past year*** that were due to a suggestion from another teacher in your school OR due to your having observed another teacher in your school.

Do NOT include changes that were due to a principal, or someone outside of your school, that you were required to make, or that occurred as a regular part of the school calendar (for example, changes that always occur when switching from fall to spring semesters).

Changes in
classroom material that you use

**Mark all
that apply**

- A. Handouts.....
- B. Books.....
- C. Hands-on learning materials.....
- D. Computer software.....
- E. Assessments (tests).....
- F. Behavior charts.....
- G. Parent communication product (for example, daily reports).....
- H. Other (please describe) _____
- I. how you teach lessons that you've taught in the past.....
- J. curriculum that involve teaching new lessons.....
- K. the homework you assign students.....
- L. how you handle behavior problems involving an individual student.....
- M. your overall approach to managing student behavior in your class.....
- N. classroom management unrelated to discipline.....
- O. your own understanding of materials/procedures that you currently use.....
- P. your own understanding of the *content* of what you teach.....
- Q. your approach to teaching specific groups of students (for example, students who are less proficient in English than they are in another language).....
- R. your approach to any aspect of extra-curricular activities that you might be involved with (for example, coaching, tutoring or helping in an after school program).....

60. In general, how often do you participate in any organized group activities or meetings involving other teachers at your school...

...that primarily focus on administrative issues, such as schedules, upcoming events, and teachers work assignments?

Number of times:

- per week
- per month
- per year

...That primarily focus on issues pertaining to student instruction/behavior?

Number of times:

- per week
- per month
- per year

Please write down the date that you filled out this survey.

Month Day Year

THANK YOU!

Appendix L. Student survey

**LIC
ASSENT FORM**

Dear Student,

You are being asked to participate in a survey that asks questions about you, your friends, and your school. It will take about 35-45 minutes to answer the questions.

This is not a test. There are no right or wrong answers and you will not be timed. The answers you give are very important. Please read each question carefully and answer it based on what you really know or do. Answer the questions truthfully and as best as you can.

The survey is confidential. Do not put your names anywhere on the survey or the answer sheet. Your answers are private.

The survey is voluntary. You do not have to complete this survey.

Thank you.

.....

Yes, I am willing to take the survey.

Print Your Name

Date

Birth date:

Month / Day / Year

Please remove this page carefully from the survey

**LESSONS IN CHARACTER STUDY**

This survey is being given as part of a study to see how well the *Lessons in Character Program* is working. This survey asks about you, the student, and what you think about yourself, your school, your teachers, and your classmates. This school year and next school year you will be asked to complete similar surveys such as this one. Each one should take about 35 minutes to complete. No one at this school will know how you answered the survey. Your answers are private.

Please do NOT write your name on the survey. You do not have to answer all of the survey questions. Please write your answers on this survey paper. Mark only one answer unless told to "Mark All That Apply."

Please read each question and fill in the bubble that goes with your answer.

Like this: Not like this:

Example 1

| Strongly Disagree | Disagree | Don't Know | Agree | Strongly Agree |
|-------------------|----------|------------|-------|----------------|
|-------------------|----------|------------|-------|----------------|

I enjoy learning Math

Example 2

How old are you: Years

According to the Paperwork Reduction Act of 1995, no persons are required to answer this survey unless it displays a valid OMB control number. The valid OMB control number for this information collection is 1850-0826. The time required to complete the survey should be about 35 minutes, including the time to review instructions. If you have any comments about the accuracy of the time estimate(s) or suggestions for improving this survey, please write to: U.S. Department of Education, Washington, D.C. 20202-4651. If you have comments or concerns about your individual submission of this form, write directly to: Rafael Valdivieso, U.S. Department of Education, 555 New Jersey Avenue, NW, Room 506E, Washington, D.C. 20208.

This collection of information is authorized by Public Law 107-279 Education Sciences Reform Act of 2002, Title I, Part C, Sec. 151(b) and Sec. 153(a). Participation is voluntary. Your responses are protected from disclosure by federal statute (PL 107-279 Title I, Part C. Sec. 183). All responses that relate to or describe identifiable characteristics of individuals may be used only for statistical purposes and may not be disclosed, or used, in identifiable form for any other purpose, unless otherwise compelled by law. Data will be combined to produce statistical reports. No individual data that links your name, address, telephone number, or identification number with your responses will be included in the statistical reports.

**LESSONS IN CHARACTER STUDY**

Fill in the bubble next to your answer. Mark only one answer for each statement. Remember, your answers are private and no one will know how you answered.

1. Are you female or male? Female
 Male

For the next two questions write your answer in the boxes.

2. How old are you?

Years

3. What is your date of birth?

MONTHDAYYEAR

Bubble in your agreement or disagreement with the following.

| | <u>Strongly Disagree</u> | <u>Disagree</u> | <u>Don't Know</u> | <u>Agree</u> | <u>Strongly Agree</u> |
|--|--------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| <u>The students in this school...</u> | | | | | |
| 4. Are nice to each other. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. Work together on things. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Make new students feel welcome. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. Get along well together. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Work out problems without fighting. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. Take good care of school property. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 10. Take responsibility for what they do. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

| | <u>Strongly Disagree</u> | <u>Disagree</u> | <u>Don't Know</u> | <u>Agree</u> | <u>Strongly Agree</u> |
|---|------------------------------|-----------------------|-----------------------|-----------------------|---------------------------|
| <u>The students in this school...</u> | | | | | |
| 11. Respect their teachers. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. Think it's important to be a good citizen. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 13. Treat one another fairly. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 14. Tell the truth. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 15. Are treated fairly by the adults in the school. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 16. Are expected to get along. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 17. Are expected to be kind and caring. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 18. Are expected to treat each other fairly. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 19. Are expected to obey the rules. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 20. Are expected to tell the truth. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 21. Are allowed to work together in classes. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 22. Plan things together with their teachers. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 23. Have a say in what goes on in their classes. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 24. Decide rules together with their teachers. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 25. Help plan what they do in school. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

LESSONS IN CHARACTER STUDY
I would...

| | <u>Strongly Disagree</u> | <u>Disagree</u> | <u>Don't Know</u> | <u>Agree</u> | <u>Strongly Agree</u> |
|---|--------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 26. Help someone learn something. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 27. Cheer up someone who is feeling sad. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 28. Get help if I saw a person being picked on. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 29. Be a friend to someone who is being teased. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 30. Get help for someone who is hurt. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Below are some things that can happen to kids every day. Think about how YOU feel, or would feel when these things happen to you. Fill in the bubble next to your answer. Fill in only one answer for each statement. Remember, your answers are private and no one will know how you answered.

| | <u>Yes</u> | <u>Sometimes</u> | <u>No</u> |
|---|-----------------------|-----------------------|-----------------------|
| 31. When I'm mean to someone, I usually feel bad about it later. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 32. I'm happy when the teacher says my friend did a good job. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 33. I understand how other kids feel. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 34. Other people's problems really bother me. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 35. I feel happy when my friend gets a good grade. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 36. When I see a kid who is upset it really bothers me. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 37. It's easy for me to tell when my mom or dad has a good day at work. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 38. It bothers me when my teacher doesn't feel well. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

| | | <u>Yes</u> | <u>Sometimes</u> | <u>No</u> |
|-----|---|-----------------------|-----------------------|-----------------------|
| 39. | I feel sorry for kids who can't find anyone to hang out with. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 40. | Seeing a kid who is crying makes me feel like crying. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 41. | When I see someone who is happy, I feel happy too. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Here are some statements about your school. For each one, fill in the bubble to tell us how much you agree or disagree with the statement. You should only mark one bubble for each statement.

| | | <u>Disagree A LOT</u> | <u>Disagree a little</u> | <u>Agree a little</u> | <u>Agree A LOT</u> |
|-----|---|---------------------------|------------------------------|---------------------------|------------------------|
| 42. | When I'm in class, I participate in class discussion. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 43. | I try hard to do well in school. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 44. | In class, I work as hard as I can. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 45. | I pay attention in class. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 46. | When I am in class, I listen very carefully. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 47. | When I'm in class, I just act like I'm working. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 48. | I don't try very hard in school. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 49. | In class, I just do enough to get by. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 50. | When I'm in class, I think about other things. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 51. | When I'm in class, my mind wanders. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Below are some things that kids do from time to time. How many times in the last two weeks has each happened to YOU? Fill in only one bubble for each statement. Remember, your answers are private and no one will see how you answered, not even your teacher.

| | <u>Never</u> | <u>Once or twice</u> | <u>A few times</u> | <u>Many times</u> |
|---|-----------------------|------------------------------|------------------------|-----------------------|
| <u>In the last two weeks...</u> | | | | |
| 52. I teased a kid at school. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 53. I was sent home from school for bad behavior. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 54. I was loud or made so much noise at school that I got in trouble. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 55. I pushed, shoved, or hit a kid from school. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 56. I took things at school without paying for them, such as food in the lunchroom. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 57. I called a kid at school a bad name. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 58. I said that I would hit a kid at school. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 59. I left out another kid on purpose. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 60. I made up something about other students to make other kids <u>not</u> like them anymore. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 61. I broke or ruined something on purpose that belonged to the school. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 62. I took something from school from the teacher or other kids that did not belong to me. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 63. I copied other students' homework or copied off the other students' tests. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 64. I skipped school or didn't go to class without an excuse. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |



Student Survey

LESSONS IN CHARACTER STUDY



Thank you for taking this survey!

Appendix M. Teacher implementation log

OMB No: 1850-0826
Expiration Date: 5/31/2010

Teacher Report

LIC Implementation Log

4th Grade

NOTICE: According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this information collection is 1850-0826. If you have any comments concerning the accuracy of the time estimate(s) or suggestions for improving this form, please write to: U.S. Department of Education, Washington, D.C. 20202-4651. If you have comments or concerns regarding the status of your individual submission of this form, write directly to: Rafael Valdivieso, U.S. Department of Education, 555 New Jersey Avenue, NW, Room 506E, Washington, D.C. 20208.

Responses to this data collection will be used only for statistical purposes. The reports prepared for this study will summarize findings across the sample and will not associate responses with a specific district or individual. We will not provide information that identifies you or your school or your district to anyone outside the study team, except as required by law.

Please contact Dr. Barbara Dietsch with any questions at 562.799.5126 or by email at bdietsc@wested.org.

LESSONS IN CHARACTER ★ LESSON TRACKING CHARTS
 4th Grade

A. Satisfaction – September- December 2008

| | Very Unsatisfied | Very Satisfied | | |
|--|--------------------------|--------------------------|--------------------------|--------------------------|
| 1. How satisfied are you with the <i>Lessons in Character</i> Program? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2. How satisfied are you with the <i>Daily Oral Language</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 3. How satisfied are you with the <i>Writing with Character</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

B. Overall Implementation – September - December 2008

| | Not at all as written | Sometimes as written | Mostly as written | Exactly as written |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| 4. In general, to what degree did you teach the lessons as written in the <i>Lessons in Character</i> Teacher Management Guide? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5. In general, to what degree did you teach the lessons as written in the <i>Daily Oral Language</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 6. In general, to what degree did you teach the lessons as written in the <i>Writing with Character</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

C. Lesson Tracking and Implementation – September - December 2008

| | Number of Lessons | | | | | |
|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 0 | 1-5 | 6-10 | 11-15 | 16-20 | 21+ |
| 7. Approximately how many lessons from <i>Lessons in Character</i> did you teach between Sept and Dec? | <input type="checkbox"/> |

| | Number of Lessons | | | | | |
|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 0 | 1-15 | 16-30 | 31-60 | 61-90 | 91+ |
| 8. Approximately how many lessons from <i>Daily Oral Language</i> did you teach between Sept and Dec? | <input type="checkbox"/> |

| | Number of Lessons | | | | | |
|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 0 | 1-5 | 6-10 | 11-15 | 16-20 | 21+ |
| 9. Approximately how many lessons from <i>Writing with Character</i> did you teach between Sept and Dec? | <input type="checkbox"/> |

Please FAX in December, just before winter break

LESSONS IN CHARACTER ★ LESSON TRACKING CHARTS
 4th Grade

A. Satisfaction – January - March 2009

| | Very Unsatisfied | Very Satisfied | | |
|--|--------------------------|--------------------------|--------------------------|--------------------------|
| 1. How satisfied are you with the <i>Lessons in Character</i> Program? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2. How satisfied are you with the <i>Daily Oral Language</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 3. How satisfied are you with the <i>Writing with Character</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

B. Overall Implementation – January - March 2009

| | Not at all as written | Sometimes as written | Mostly as written | Exactly as written |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| 4. In general, to what degree did you teach the lessons as written in the <i>Lessons in Character</i> Teacher Management Guide? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5. In general, to what degree did you teach the lessons as written in the <i>Daily Oral Language</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 6. In general, to what degree did you teach the lessons as written in the <i>Writing with Character</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

C. Lesson Tracking and Implementation – January - March 2009

| | Number of Lessons | | | | | |
|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 0 | 1-5 | 6-10 | 11-15 | 16-20 | 21+ |
| 7. Approximately how many lessons from <i>Lessons in Character</i> did you teach <u>between Jan and Mar 2009</u> ? | <input type="checkbox"/> |
| | Number of Lessons | | | | | |
| | 0 | 1-15 | 16-30 | 31-60 | 61-90 | 91+ |
| 8. Approximately how many lessons from <i>Daily Oral Language</i> did you teach <u>between Jan and Mar 2009</u> ? | <input type="checkbox"/> |
| | Number of Lessons | | | | | |
| | 0 | 1-5 | 6-10 | 11-15 | 16-20 | 21+ |
| 9. Approximately how many lessons from <i>Writing with Character</i> did you teach <u>between Jan and Mar 2009</u> ? | <input type="checkbox"/> |

Please FAX in March, or just before spring break

LESSONS IN CHARACTER ★ LESSON TRACKING CHARTS
 4th Grade

A. Satisfaction – April - May 2009

| | Very Unsatisfied | | | | Very Satisfied |
|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 1. How satisfied are you with the <i>Lessons in Character</i> Program? | <input type="checkbox"/> |
| 2. How satisfied are you with the <i>Daily Oral Language</i> materials? | <input type="checkbox"/> |
| 3. How satisfied are you with the <i>Writing with Character</i> materials? | <input type="checkbox"/> |

B. Overall Implementation – April - May 2009

| | Not at all as written | Sometimes as written | Mostly as written | Exactly as written |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| 4. In general, to what degree did you teach the lessons as written in the <i>Lessons in Character</i> Teacher Management Guide? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5. In general, to what degree did you teach the lessons as written in the <i>Daily Oral Language</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 6. In general, to what degree did you teach the lessons as written in the <i>Writing with Character</i> materials? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

C. Lesson Tracking and Implementation – April - May 2009

| | Number of Lessons | | | | | |
|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | 0 | 1-5 | 6-10 | 11-15 | 16-20 | 21+ |
| 7. Approximately how many lessons from <i>Lessons in Character</i> did you teach between April-May 2009? | <input type="checkbox"/> |
| | Number of Lessons | | | | | |
| | 0 | 1-15 | 16-30 | 31-60 | 61-90 | 91+ |
| 8. Approximately how many lessons from <i>Daily Oral Language</i> did you teach between April-May 2009? | <input type="checkbox"/> |
| | Number of Lessons | | | | | |
| | 0 | 1-5 | 6-10 | 11-15 | 16-20 | 21+ |
| 9. Approximately how many lessons from <i>Writing with Character</i> did you teach between April-May 2009? | <input type="checkbox"/> |

Please FAX in May or June, just before the end of the school year

LESSONS IN CHARACTER ★ LESSON TRACKING CHARTS
 4th Grade

Please note the date you taught each of the lessons/activities listed. Use an "X" to note the degree to which you followed the "Teaching Strategies" for each lesson.

| | Student Workbook Pages | Date | Degree of Implementation | | | |
|--|------------------------|------|--------------------------|--------------------------|--------------------------|--------------------------|
| | | | All | Most | Some | None |
| STAR: A Problem-solving Process | NA | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| TRUSTWORTHINESS: Be a Courageous Person | | | | | | <input type="checkbox"/> |
| Lesson 1: Defining the Theme | | | | | | |
| Activity 1: My Journal | 1 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 2: "The Honest Thumb" | 2 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Group activity – write and perform skits | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a courageous person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 2: In School | | | | | | |
| Activity 3: Case Study: The big red "F" | 3-4 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 4: Showing courage at school | 5 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a courageous person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 3: Personal commitment | | | | | | |
| Activity 5: My Journal | 6 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 6: "Holding up the Sky" | 7 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 7: My positive self-talk | 8 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a courageous person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 4: In the Community | | | | | | |
| Activity 8: "I Never Before Heard of Such a Thing" | 9 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 9: A community leader with courage | 10 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a courageous person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Comments: | | | | | | |

LESSONS IN CHARACTER ★ LESSON TRACKING CHARTS
 4th Grade

| | Student Workbook Pages | Date | Degree of Implementation | | | |
|---|------------------------------|------|--------------------------|--------------------------|--------------------------|--------------------------|
| | | | All | Most | Some | None |
| RESPECT: Be a Considerate Person | | | | | | |
| Lesson 1: Defining the Theme | | | | | | |
| Activity 10: "Age and Wisdom" | 11 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 11: Considerate cereal | 12 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a considerate person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 2: In School | | | | | | |
| Review STAR Model | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 12: Case Study: "The Monster" | 13-14 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 13: Being more considerate at school | 15 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a considerate person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 3: Personal commitment | | | | | | |
| Activity 14: I feel... | 16 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 15: "The Mouse, the Bird, and Sausage" | 17 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 15: My Journal | 18 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a considerate person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 4: In the Community | | | | | | |
| Activity 16: "The Lion's Whisker" | 19 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Review STAR Model | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 17: Rules of consideration | 20 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Writing activity: "The Selfish Giant" | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Comments: | | | | | | |

LESSONS IN CHARACTER ★ LESSON TRACKING CHARTS
 4th Grade

| | Student Workbook Pages | Date | Degree of Implementation | | | |
|--|------------------------------|------|--------------------------|--------------------------|--------------------------|--------------------------|
| | | | All | Most | Some | None |
| RESPONSIBILITY Be a Reliable Worker | | | | | | |
| Lesson 1: Defining the Theme | | | | | | |
| Activity 18: "The Lazy Tunrit" | 21 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 19: My Journal | 22 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a reliable worker poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 2: In School | | | | | | |
| Activity 20: What if....? | 23 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 21: Case Study: "Partners" | 24-25 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 22: You can rely on me | 26 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a reliable worker poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 3: Personal commitment | | | | | | |
| Review STAR Model | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 23: "The Wasp and the Bee" | 27 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 24: "Truth and Falsehood" | 28 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 25: Relying on STAR | 29 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a reliable worker poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 4: In the Community | | | | | | |
| Activity 26: Am I reliable at home? | 30 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 27: "The Great White Condor" | 31 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Group activity: Write ad for job | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a reliable worker poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Comments: | | | | | | |

LESSONS IN CHARACTER ★ LESSON TRACKING CHARTS
 4th Grade

| | Student Workbook Pages | Date | Degree of Implementation | | | |
|--|------------------------------|------|--------------------------|--------------------------|--------------------------|-------------------------------------|
| | | | All | Most | Some | None |
| FAIRNESS Be a Good Sport | | | | | | <input checked="" type="checkbox"/> |
| Lesson 1: Defining the Theme | | | | | | |
| Activity 28: "The Carvers" | 32 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 29: My Journal | 33 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a good sport poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 2: In School | | | | | | |
| Group activity: Create Discussion Rules | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 30: Case Study: "So many things to learn" | 34-35 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Review STAR Model | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 31: No put-downs allowed | 36 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a good sport poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 3: Personal commitment | | | | | | |
| Activity 32: My Journal | 37 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 33: "The Impatient Mule Drivers" | 38 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 34: You be the leader | 39 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a good sport poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 4: In the Community | | | | | | |
| Activity 35: "The Ruler of the Trees" | 40 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Group activity: Create laws for the community | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 36: Just the facts, please | 41 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a good sport poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Comments: | | | | | | |

LESSONS IN CHARACTER ★ LESSON TRACKING CHARTS
 4th Grade

| | Student Workbook Pages | Date | Degree of Implementation | | | |
|--|------------------------------|------|--------------------------|--------------------------|--------------------------|--------------------------|
| | | | All | Most | Some | None |
| CARING Be a kind person | | | | | | |
| Lesson 1: Defining the Theme | | | | | | |
| Activity 37: My Journal | 42 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 38: "The Guest on New Year's Eve" | 43 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 39: How kind am I? | 44 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: List of kind actions | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 2: In School | | | | | | |
| Activity 40: Case Study: "The new kid at school" | 45-46 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Write and deliver notes of appreciation | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a kind person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 3: Personal commitment | | | | | | |
| Activity 41: My Journal | 47 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 42: "Ask a Foolish Question" | 48 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Review STAR Model | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 43: Be prepared | 49 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a kind person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 4: In the Community | | | | | | |
| Activity 44: "The House of Peace" | 50 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 45: Be kind to a family member | 51 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 46: Thanks for your kindness | 52 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a kind person poster | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Comments: | | | | | | |

LESSONS IN CHARACTER ★ LESSON TRACKING CHARTS
 4th Grade

| | Student Workbook Pages | Date | Degree of Implementation | | | |
|--|------------------------------|------|--------------------------|--------------------------|--------------------------|--------------------------|
| | | | All | Most | Some | None |
| CITIZENSHIP Be a protector of the environment | | | | | | |
| Lesson 1: Defining the Theme | | | | | | |
| Activity 47: "Hurt no Living Thing" | 53 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 48: The three Rs | 54 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Recycled products | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 2: In School | | | | | | |
| Activity 49: Case Study: Start in your own backyard | 55-57 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Create "Environmental Bill of Responsibilities" | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a protector of the environment | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 3: Personal commitment | | | | | | |
| Activity 50: "Birdfoot's Grampa" | 58 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 51: My Journal | 59 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a protector of the environment | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Lesson 4: In the Community | | | | | | |
| Activity 52: "Lady Cápa and the Eagle" | 60 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Activity 53: A letter of request | 61 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Class Project: Be a protector of the environment | | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Be a STAR | 62 | | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Comments: | | | | | | |

References

Aber, J. L., Brown, J. L., Jones, S. M., Berg, J., and Torrente, C. (2011). School-based strategies to prevent violence, trauma, and psychopathology: The challenge of going to scale. *Development and Psychopathology*, 23:411–421

Aspen Declaration on Character Education. (1992). Retrieved December 1, 2010, from <http://charactercounts.org/overview/aspen.html>.

Battistich, V. (2005). *Character education, prevention, and positive youth development*. Washington, DC: Character Education Partnership. Retrieved December 6, 2010, from http://www.character.org/uploads/PDFs/White_Papers/White_Paper_Battistich.pdf.

Beets, M. W., Flay, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Li, K.-K., Burns, K., Washburn, I. J., and Durlak, J. A. (2009). Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary students in Hawai'i. *American Journal of Public Health*, 99(8):1438-144.

Benninga, J. S., Berkowitz, M. W., Kuehn, P., and Smith, K. (2003). The relationship between character education and academic achievement. *Journal of Research in Character Education*, 1(1), 19–32.

Berkowitz, M. W., and Bier, M. C. (2005, February). *What works in character education: research-based character education*. Washington, DC: Character Education Partnership.

Berkowitz, M.W., and Bier, M.C. (2007). What works in character education. *Journal of Research in Character Education*, 5(1): 29–48.

Bloom, H.S. (1995) Minimum detectable effects: simple way to report the statistical power of experimental designs. *Evaluation Review*, 19(5): 547–556.

Bowen, N. K., and Bowen, G. L. (1999). Effects of crime and violence in neighborhoods and schools on the school behavior and performance of adolescents. *Journal of Adolescent Research*, 14(3), 219–341.

Brooks, B.D. (1996). *Young people's Lessons in Character: teacher management guide*. San Diego, CA: Young People's Press.

California Department of Education. (2009). *Academic performance index data file, 2007 base*. Retrieved March 13, 2009, from <http://www.cde.ca.gov/ta/ac/ap/apidatafiles.asp>.

California Department of Education. (2010a). *California standards test: technical report, spring 2009 administration*. Retrieved July 13, 2010, from <http://www.cde.ca.gov/ta/tg/sr/documents/csttechrpt09.pdf>.

California Department of Education. (2010b). *State report: growth API, 2008 growth academic performance index (API) report*. Retrieved Dec 6, 2010, from <http://api.cde.ca.gov/AcntRpt2008/2008GrthStAPI.aspx?allcds=0>.

California Department of Education. (2010c). *2009-2010 academic performance index reports: information guide*. Retrieved Dec 6, 2010, from <http://www.cde.ca.gov/ta/ac/ap/documents/infoguide09.pdf>.

California Department of Education. (2011). *Academic performance index (API)*. Retrieved August 25, 2011, from <http://www.cde.ca.gov/ta/ac/ap/>

Character Education Partnership. (2010). *State character education legislation*. Retrieved Dec 6, 2010, from http://www.character.org/uploads/PDFs/State_Character_Education_Legislation_2009-9-28.pdf.

Characterplus. (2002). *Evaluation resource guide: tools and strategies for evaluating a character education program*. St. Louis, MO: Characterplus.

Dietsch, B., Bayha J., and Zheng, H. (2005, April). *Short-term effects of a literature-based character education program among fourth grade students*. Paper presented at the 2005 Meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Donner, A.N., and Klar, N. (2000). *Design and analysis of cluster randomization trials in health research*. London: Arnold.

Flay, B. R, Berkewitz, M. W., and Bier, M. C. (2009). Elementary school-based programs theorized to support social development, prevent violence, and promote positive school climate: Description and hypothesized mechanisms of change. *Journal of Research in Character Education*, 7(2): 21-49.

Fox, J. (2002). *Bootstrapping regression models. Appendix to an R and S-PLUS companion to applied regression*. Retrieved Aug 30, 2011, from <http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-bootstrapping.pdf>.

Funk, J., Elliott, R., Bechtoldt, H., and Pasold, T. (2003). The attitudes toward violence scale: child version. *Journal of Interpersonal Violence*, 18(2), 186–196.

Furrer, C., and Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 95, 148–162.

Glanzer, P.L., and Milson, A.J. (2006). Legislating the good: a survey and evaluation of character education laws in the United States. *Educational Policy*, 20(3), 525–550.

Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Oxford University Press.

Gresham, F.M., and Elliott, S.N. (1990). *Social Skills Rating System (SSRS)*. Bloomington, MN: Pearson Assessments.

Hanson, T. L., Austin, G. A., and Lee-Bayha, J. (2004). *Ensuring that no child is left behind: how are student health risks and resilience related to the academic progress of schools*. San Francisco: WestEd.

Hedges, L.V., and Hedberg, E.C. (2007). Intraclass correlation values for planning group randomized trials in education. *Education Evaluation and Policy Analysis*, 29(1): 60–87.

Hill, C.J., Bloom, H.S., Black, A.R., and Lipsey, M.W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2(3), 172–177.

Hoskisson, K and Garrison J. (1998). Predictive reading and scientific enquiry. *Reading*, 22(2), 118-125.

Jolliffe, D., and Farrington, D.P. (2004). Empathy and offending: a systematic review and meta-analysis. *Aggressive Violent Behavior*, 9, 441–476.

Jones, S. M., Brown, J. L., and Aber, J. L., (2011). Two-year impacts of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development*, 82(2):533-554.

Kisker, E, Kalb, L, Miller, M., Sprachman, S., Carey, N, Schochet, P, and James-Burdumy, S. (2004). *Social and character development research program evaluation: supporting statement for request for OMB approval of SACD evaluation*. Princeton, NJ: Mathematica Policy Research.

Lochman, J. E., Lampron, L. B., Gemmer, T. C., and Harris, S. R. (1987). Anger coping intervention with aggressive children: a guide to implementation in school settings. In P. A. Keller and S. R. Heyman (Eds.), *Innovations in clinical practice: a source book*, 6, 339–356.

Orpinas, P. and Frankowski, R. (2001).The aggression scale: a self-report measure of aggressive behavior for young adolescents. *Journal of Early Adolescence*, 21(1), 50–67.

Orr, L, Feins, J.D, Jacob, R., Beecroft, E, Sanbonmatsu, L, Katz, L, Liebman, J.B., Kling, J.R. (2003).*Moving to Opportunity interim impacts evaluation: final report*. U.S. Department of Housing and Urban Development, Office of Policy Development and Research. Washington, DC.

Murray, D.M. (1998). *Design and analysis of group randomized trials*. New York: Oxford University Press.

Murray, D.M., and Blitstein, J.L.2003. Methods to reduce the impact of intraclass correlation in group randomized trials. *Evaluation Review*, 27(1):79–103.

Murray, D. M., and Hannan, P. J. (1990). Planning for appropriate analysis in school-based drug use prevention study. *Journal Consulting and Clinical Psychology*, 58, 458–468.

Murray, D.M., and Short, B.J. (1996).Intraclass correlation among measures related to alcohol use by school aged adolescents: estimates, correlates, and applications in intervention studies. *Journal of Drug Education*, 26(3):207–230.

Rabe-Hesketh, S., and Skrondal , A. (2008). *Multilevel and longitudinal modeling using Stata*. 2nd ed. College Station, TX: Stata Press.

Raudenbush, S. W. (1997). Statistical analysis and optimal design in cluster randomized trials. *Psychological Methods*, 2(2), 173–185.

Raudenbush, S.W., and Bryk, A.S. (2002). *Hierarchical linear models: applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.

Ryan, K. (2003, January). Character education: our high schools' missing link. *Education Week on the Web*, 22(20), 35–48.

Schafer, J.L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Scheier, L.M., Griffin, K.W., Doyle, M.M., and Botvin, G.J. (2002). Estimates of intragroup dependence for drug use and skill measures in school-based drug abuse prevention trials: an empirical study of three independent samples. *Health, Education, and Behavior*, 29(1): 85–103.

Schochet, P.Z. (2005). *Statistical power for random assignment evaluations of education programs*. Princeton, NJ: Mathematica Policy Research.

Schochet, P.Z. (2008). *Guidelines for multiple testing in experimental evaluations of educational interventions*. Princeton, NJ: Mathematical Policy Research.

Singer, J.D., and Willett, J.B., (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. New York, NY: Oxford University Press.

Snyder, F. J., Flay, B. R., Vuchinich, S., Acock, A., Washburn, I. J., Beets, M. W., and Li, K. K. (2010). Impact of the Positive Action program on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes: A matched-pair, cluster randomized, controlled trial. *Journal of Research on Educational Effectiveness*, 3(1): 26–55.

Social and Character Development Research Consortium. (2010). *Efficacy of schoolwide programs to promote social and character development and reduce problem behavior in elementary school children*. (NCER 2011-2001). Washington DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

StataCorp. 2007. *Stata Statistical Software: Release 10*. College Station, TX: StataCorp LP.

U.S. Department of Education, National Center for Educational Statistics. (2009). Common Core of Data. Public Elementary/Secondary School Universe Survey, 2008/09. Retrieved August 3, 2010, from <http://nces.ed.gov/ccd/pubschuniv.asp>.

U.S. Department of Education. (2010). *Partnerships in character education program*. Retrieved Dec 6, 2010, from <http://www2.ed.gov/programs/chartered/funding.html>.

von Hippel, P.T. (2007). Regression with missing Ys: an improved strategy for analyzing multiply-imputed data. *Sociological Methodology*, 37: 83–117.

Warden, D., and Mackinnon, S. (2003). Prosocial children, bullies and victims: an investigation of their sociometric status, empathy and social competence. *British Journal of Developmental Psychology*, 21: 367–385.

Westfall, P.H., and Young, S.S. (1993). *Resampling-based multiple testing: examples and methods for p-value adjustment*. New York: John Wiley and Sons.

What Works Clearinghouse. (2007). *Character education topic report*. Retrieved February 2, 2009, from http://ies.ed.gov/ncee/wwc/pdf/CE_TR_06_04_07.pdf.

What Works Clearinghouse. (2008). *Procedures and standards handbook (version 2.0)*. Retrieved November 23, 2010, from http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf.

Williams, M. (2000). Models of character education: perspectives and developmental issues. *Journal of Humanistic Counseling, Education & Development*, 39(1), 32–41.

Wolf, P, Gutmann, B, Puma, M., Rizzo, L., Eissa, N, and Silverberg, M. (2007). *Evaluation of the DC Opportunity Scholarship Program: impacts after one year*. U.S. Department of Education, Institute of Education Sciences. Washington, DC: U.S Government Printing Office.

